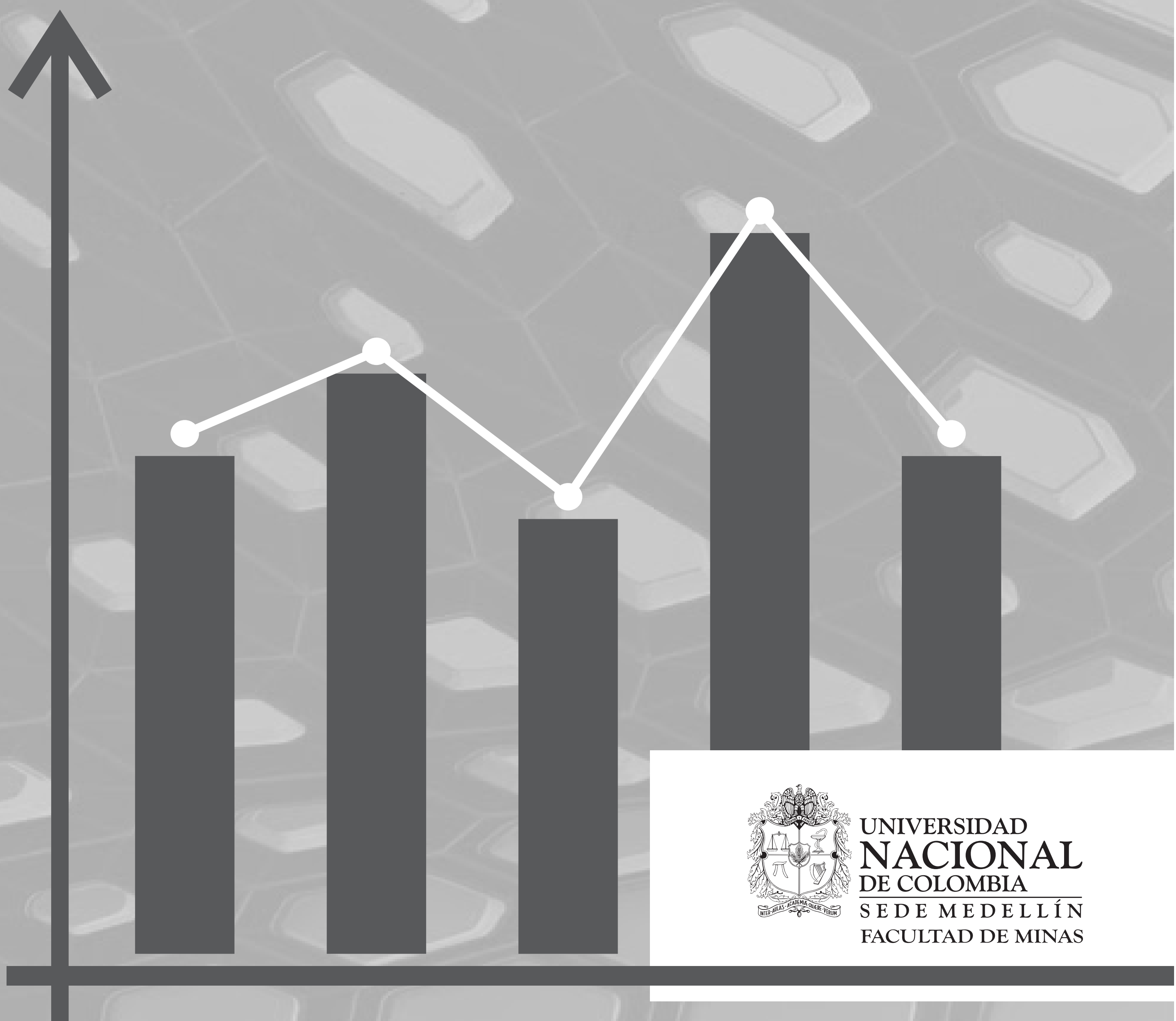
FUNDAMENTOS ESTADÍSTICOS PARA EL CONTROL ESTADÍSTICO DE CALIDAD

Hernán Darío Rendón C.



UNIVERSIDAD
NACIONAL
DE COLOMBIA
SEDE MEDELLÍN
FACULTAD DE MINAS

FUNDAMENTOS ESTADÍSTICOS PARA EL CONTROL ESTADÍSTICO DE CALIDAD

FUNDAMENTOS ESTADÍSTICOS PARA EL CONTROL ESTADÍSTICO DE CALIDAD

Hernán Darío Rendón C.

UNIVERSIDAD NACIONAL DE COLOMBIA
SEDE MEDELLÍN
FACULTAD DE MINAS

Medellín, Colombia, 2013

658.562
R35 Rendón Castaño, Hernán Darío
Fundamentos estadísticos para el control estadístico de calidad / Hernán Darío Rendón C. -- Medellín : Universidad Nacional de Colombia. Facultad de Minas, 2013.
88 p.

ISBN :

1. CONTROL DE CALIDAD. 2. MUESTREO (ESTADISTICA). 3. PROBABILIDADES. 4. VARIABLES ALEATORIAS. 5. ANALISIS DE VARIANZA. I. Tít.

Catalogación en la publicación Universidad Nacional de Colombia

Fundamentos Estadísticos para el Control Estadístico de Calidad
Hernán Darío Rendón C.
Universidad Nacional de Colombia, Sede Medellín, Facultad de Minas

Colección Facultad de Minas

Primera edición: Medellín, octubre de 2013
ISBN: 978-958-761-630-9

Caratula: Imágenes diseñadas por Laura Rendón G.

Coordinación editorial
Centro Editorial de la Facultad de Minas
Universidad Nacional de Colombia, Sede Medellín
Carrera 80 No. 65 – 223, Bloque M9-103
Teléfono: (57-4) 425 53 43 ceditorial_med@unal.edu.co

Impreso y hecho en Medellín, Colombia

CONTENIDO

MODULO 4 – DISTRIBUCIONES DISCRETAS DE PROBABILIDAD

MODULO 5 – LA DISTRIBUCIÓN NORMAL

MODULO 6 – TEORIA DEL MUESTREO

DEDICATORIA

A mi hija: LAURA, perenne compañera de viaje.

Hernán Darío.

“No por tener la cabeza en un congelador y los pies en un horno, se disfruta de una agradable temperatura en el ombligo”

Anónimo

INTRODUCCIÓN

En mis años de experiencia docente, en el área de Control Estadístico de la Calidad, he podido observar, que en términos generales, los estudiantes abordan la asignatura con bajos niveles de suficiencia en la comprensión y limitada capacidad para la aplicación de las herramientas estadísticas que sustentan este campo de la calidad.

Por lo anterior, el desarrollo exitoso de la asignatura Control Estadístico de Calidad, requiere necesariamente la inclusión del “repaso” de contenidos estadísticos que le permitan al estudiante lograr mayores niveles de comprensión, contextualización y habilidad para la aplicación y uso de estas herramientas.

El material académico contenido en este texto, pretende ofrecer al estudiante los contenidos conceptuales y prácticos de la estadística en que se fundamenta el control de la calidad para los procesos productivos de bienes y servicios.

El diseño de este material permite al estudiante abordar estos contenidos de manera independiente, a partir de seis módulos integrados por conceptos teóricos, ejemplos prácticos, talleres de aplicación y referencias bibliográficas.

En relación con los conceptos teóricos de cada uno de los módulos, es pertinente observar, que estos se limitan a las definiciones básicas. La consulta de las referencias bibliográficas es imprescindible para obtener mayores niveles de detalle y profundidad conceptual.

La aplicación de los conceptos se ilustra a partir de ejemplos relacionados con situaciones propias del área de calidad en las actividades productivas de bienes y servicios. Además, de su solución y análisis se incluyen los objetivos pretendidos de su presentación.

Cada taller incluye una selección de problemas tipo relacionados con la temática de cada uno de los módulos. Es responsabilidad del estudiante abordar la solución e interpretación de los resultados de los problemas.

El primer módulo se ocupa del tema de los Espacios del Muestreo y Sucesos propios de los experimentos. Los conceptos de Probabilidad son el tema del modulo 2. El módulo 3 aborda los conceptos de: Variable Aleatoria y Distribuciones de Probabilidad. Los módulos 4 y 5 tratan las distribuciones de probabilidad discretas y continuas más usadas en el control estadístico de calidad. El módulo 6 examina los conceptos básicos y aplicaciones de la Teoría del Muestreo. Anexo, se presenta la solución de los problemas de los talleres planteados en cada uno de los módulos.

Se incluye, además, una invitación al estudiante para que comparta este material, y una serie de recomendaciones para lograr que su uso fortalezca el proceso de enseñanza – aprendizaje en el campo que nos ocupa.

Finalmente, deseo agradecer las enseñanzas recibidas de mis colegas, y en especial, de los estudiantes, sin sus aportes este material no hubiese sido posible.

El Autor

2003

AL ESTUDIANTE

El presente documento está dirigido a los estudiantes del curso de Control Estadístico de Calidad, asignatura del núcleo profesional del programa curricular de Ingeniería Industrial y ofrecida por la Escuela de Ingeniería de la Organización de la Facultad de Minas de la Universidad Nacional de Colombia Sede de Medellín.

Los objetivos de este material se pueden resumir en los siguientes puntos:

- Refrescar los conceptos estadísticos que fundamentan la práctica del Control Estadístico de Calidad en la actividad productiva de bienes y servicios.

- Contextualizar las técnicas estadísticas y desarrollar habilidades para el uso inteligente de estas herramientas en el desempeño profesional del futuro ingeniero en esta área.

- Propiciar un mayor protagonismo y autonomía del estudiante en el proceso enseñanza – aprendizaje

El logro de los citados objetivos, dependerá en gran medida del trabajo académico autónomo del estudiante y el correcto uso de este material.

Se sugiere abordar los módulos del texto de manera secuencial hasta lograr su total comprensión e idoneidad en la solución de los problemas planteados en los talleres. Evite consultar el anexo de respuestas antes de intentar la solución e interpretación de los problemas.

Auto evalúe continuamente los niveles de comprensión y las capacidades de aplicación de estas herramientas en la solución de problemas propios de este campo de la ingeniería.

MODULO 1: ESPACIOS DE MUESTRAS Y SUCESOS

OBJETIVOS DEL MODULO:

- Comprender el concepto de experimento, sus condiciones y características.
- Comprender los conceptos de espacio de muestras y sucesos o eventos de un experimento.
- Aprender reglas para determinar el número de elementos del espacio de muestras de un experimento (Combinaciones y permutaciones)

CONTENIDO:

PRERREQUISITOS:

- Teoría de conjuntos
- Operaciones de conjuntos
- Estadística descriptiva

1. **EXPERIMENTO:** Un proceso que genera resultados observables. Son de interés sólo aquellos experimentos que tengan las siguientes características:

   - **Pueden repetirse indefinidamente sin cambiar sus condiciones:** Esta característica significa que los resultados del experimento en cada una de sus repeticiones estén sometidos a las mismas condiciones. Suponga que se desea registrar los kilogramos de pesca por día de una cooperativa pesquera. Evidentemente los resultados de este ”experimento” variarán dependiendo de la época del año en que éste se realice. Por lo tanto, las condiciones serán diferentes en época de subienda, por ejemplo.

Considere el experimento de extraer de una caja que contiene 2 bolas rojas y 2 bolas blancas, 2 bolas al azar. Observe que si el experimento se realiza dos veces y luego de su primera repetición no se restituyen las bolas extraídas a la caja, las condiciones de la segunda repetición no son comparables con las de la primera.

- **Aunque no es posible saber el resultado particular del experimento si se conoce el conjunto de todos los posibles resultados:** Esta característica se cumple en múltiples experimentos, considere, por ejemplo, el experimento de lanzar un dado, pese a no saber en que número caerá en cualquiera de sus repeticiones, con seguridad caerá en 1, 2, 3, 4, 5 o 6.

- **Al repetir el experimento un gran número de veces, los resultados obtenidos presentan un patrón de regularidad:** Esta condición está íntimamente ligada con la primera condición, observe, por ejemplo, que al lanzar una moneda equilibrada un gran número de veces, es de esperarse que la mitad de las veces resulte cara.

2. **ESPACIO DE MUESTRAS:** Conjunto de todos los resultados posibles de un experimento. Se denota como el conjunto S y representa el conjunto universal de los resultados del experimento. La tabla 1.1. muestra algunos ejemplos:

EXPERIMENTO	ESPACIO DE MUESTRAS
Lanzar dos dados (suma puntos)	S = {2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}
Extraer 2 artículos de un lote que contiene algunos defectuosos (número de defectuosos extraídos)	S = {0, 1, 2}
Extraer 4 artículos de un lote que contiene algunos defectuosos (sucesión de buenos (b) y defectuosos (m) extraídos)	S = {bbbb, bbbm, bbmb, bmbb, mbbb, bbmm, mmbb, bmbm, mbmb, mbbm, bmmb, bmmm, mbmm, mmbm, mmmb, mmmm}
Examinar las piezas negras de un juego de ajedrez (contar las defectuosas)	S = {0, 1, 2, 3, ..., 14, 15, 16}
Registrar continuamente la temperatura del agua de una piscina durante un período de 12 horas	S = {f / f es una función continua y 30º ≤ f(t) ≤ 42º para todo t} (Los valores de 30º y 42º se especifican como los valores mínimo y máximo posibles)
Registrar la estatura de los habitantes de una población	S = {e / e es una función continua y 0.25 m ≤ f(e) ≤ 2.1 m para todo e} (Los valores de 0.25 y 2.10 se especifican como los valores mínimo y máximo posibles)

Tabla 1.1. - Espacios de muestras

Ejemplo 1: Considere el experimento de extraer de una bolsa, al azar, una ficha de dominó y registrar la ficha obtenida.

- Determine el espacio de muestras asociado con este experimento

Objetivo del ejemplo 1:

- Mostrar que la situación presentada si constituye un experimento
- Mostrar cómo determinar el espacio de muestras asociado al experimento

Solución:

- La extracción de la ficha al azar, remplazándola luego de cada extracción, puede repetirse indefinidamente, sin cambiar las condiciones.
- A pesar de no saberse el resultado (ficha obtenida) en una extracción, si se sabe que este resultado será una de las 28 fichas del dominó.

Las fichas del dominó, son:

FICHA			**FICHA**	
0	0		2	3
0	1		2	4
0	2		2	5
0	3		2	6
0	4		3	3
0	5		3	4
0	6		3	5
1	1		3	6
1	2		4	4
1	3		4	5
1	4		4	6
1	5		5	5
1	6		5	6
2	2		6	6

Dado que el resultado de interés es la ficha obtenida al hacer la extracción, se tiene que el espacio de muestras asociado, es: S = { (0-0), (0-1), (0-2), (0-3), (0-4), (0-5), (0-6), (1-1), (1-2), (1-3), (1-4), (1-5), (1-6), (2-2), (2-3), (2-4), (2-5), (2-6), (3-3), (3-4), (3-5), (3-6), (4-4), (4-5), (4-6), (5-5), (5-6), (6-6)}

3. **SUCESO:** El suceso A respecto al espacio de muestras S asociado con un experimento es un conjunto de resultados posibles. El suceso A es un subconjunto del espacio de muestras S.

Se dice que los sucesos A y B, asociados a un experimento, son mutuamente excluyentes si no pueden ocurrir simultáneamente en una repetición del experimento. Considere el lanzamiento de una moneda, si se define el suceso C (cae cara) y el suceso S (cae sello), C y S son mutuamente excluyentes, puesto que la moneda caerá o bien cara, o bien sello. Esto equivale a que la intersección de los conjuntos C y S es el conjunto vacío: $C \cap S = \emptyset$.

Los sucesos A y B son independientes si la ocurrencia del suceso A en una repetición del experimento no tiene ninguna influencia en la ocurrencia del suceso B en otra repetición del experimento.

Ejemplo 2: Determine los siguientes sucesos en relación con el experimento del ejemplo 1:

- A_1 = Sacar la ficha blanca, blanca
- A_2 = Sacar una ficha doble
- A_3 = Sacar una ficha cuyo número total de puntos sea 4.
- A_4 = Sacar la ficha blanca, ocho
- A_5 = Sacar cualquier ficha

Objetivo del ejemplo 2: Determinar los subconjuntos A_K del espacio de muestras S del experimento, asociados con los sucesos de interés.

Solución:

- $A_1 = \{(0-0)\}$
- $A_2 = \{(0-0), (1-1), (2-2), (3-3), (4-4), (5-5), (6-6)\}$
- $A_3 = \{(0-4), (1-3), (2-2)\}$
- $A_4 = \{\Phi\}$ Conjunto vacío
- $A_5 = \{ (0-0), (0-1), (0-2), (0-3), (0-4), (0-5), (0-6), (1-1), (1-2), (1-3), (1-4), (1-5), (1-6), (2-2), (2-3), (2-4), (2-5), (2-6), (3-3), (3-4), (3-5), (3-6), (4-4), (4-5), (4-6), (5-5), (5-6), (6-6)\}$

4. NUMERO DE RESULTADOS DEL ESPACIO DE MUESTRAS DE UN EXPERIMENTO:

Los espacios de muestras pueden ser: Finitos, Infinitos numerables o Infinitos no numerables. Así, el espacio de muestras del experimento de lanzar un dado es finito (6 resultados posibles).
El espacio de muestras de la generación de una lista de los números primos es infinito numerable, Mientras que el espacio de muestras asociado con el experimento de medir la resistencia de una barra de acero es infinito no numerable.

Para el caso de espacios de muestras finitos es útil en muchas situaciones determinar el número de resultados posibles.

Retomando el ejemplo 1, se observa que el experimento puede ocurrir de 28 maneras diferentes, esto es, el número de elementos del espacio de muestras S = 28.

Como ilustración adicional, considere un juego de lotería de cuatro cifras mediante el giro de cuatro ruedas numeradas del 0 al 9, es evidente que el espacio de muestras es S = {0000, 0001, 0002, ..., 9998, 9999}, y por lo tanto es espacio de muestras está constituido por 10000 resultados.

En muchos casos la determinación del número de elementos del espacio de muestras requiere el uso de métodos especiales, los cuales se exponen a continuación.

- **REGLA DE MULTIPLICACIÓN:** Si una operación puede realizarse de n_1 maneras y por cada una de estas, una segunda operación puede realizarse de n_2 maneras, luego, la realización de ambas operaciones puede hacerse de $(n_1)x(n_2)$ maneras.

  Ejemplo 3: Considere el experimento de extraer al azar de manera consecutiva 2 balotas (sin reemplazo) de una urna que tiene 10 balotas numeradas del 0 al 9. ¿Cuál es el número posible de resultados de este experimento?

  Objetivo del ejemplo 3: Mostrar la aplicación del principio de multiplicación para determinar el número de resultados del experimento.

Solución:

Observe que la primera balota extraída puede resultar de 10 maneras diferentes, la segunda balota puede resultar de 9 maneras, dado que ya había sido extraída una en la primera extracción. Por lo tanto, el número total de maneras estará dado por el producto 10 x 9 = 90.

El anterior principio puede extenderse a k operaciones donde la primera operación puede realizarse de n_1 maneras y por cada una de estas una segunda operación puede realizarse de n_2 maneras y por cada una de las primeras dos operaciones una tercera operación puede realizarse de n_3 maneras y así sucesivamente, las k operaciones pueden realizarse de $(n_1) \times (n_2) \times (n_3) \times \ldots\ldots \times (n_k)$ maneras.

Ejemplo 4: La producción de un articulo requiere de tres materias primas. La materia prima 1 puede comprarse a tres proveedores, la materia prima 2 a cinco proveedores y la materia prima 3 a cuatro proveedores, De cuántas maneras pueden comprarse las tres materias primas?

Objetivo del ejemplo 4: Mostrar la aplicación de la regla de multiplicación para determinar el número total de resultados de un experimento.

Solución:

La provisión de las tres materias primas puede hacerse de 3 x 5 x 4 = 60 maneras.

- **PERMUTACIONES:** Una permutación es una selección o arreglo de todos o parte de los objetos de conjunto.

  El número de permutaciones de n objetos tomando n a la vez, denotado como ${}_nP_n = n!$

  Ejemplo 5: Una compañía de ingenieros requiere la asignación de tres residentes para tres obras (1 residente para cada obra). De cuántas maneras pueden asignarse?

  Objetivo del ejemplo 5: Mostrar la aplicación de las permutaciones para determinar el número de resultados posibles de un experimento.

Solución: Observe que el primer residente puede asignarse de 3 maneras, el segundo residente de 2 maneras (dado que una de las obras ya tiene asignación) y el tercero de 1 manera (dado que dos obras ya han sido asignadas). Aplicando el principio de multiplicación, el número total de posibles asignaciones es de 3 x 2 x 1 = 3! = 6 maneras.

El número de permutaciones de n objetos tomando m de estos objetos (donde $m \leq n$) , es: ${}_nP_m = n! / (n - m)!$

Ejemplo 6: Suponga que la compañía dispone de 4 residentes para asignarlos a 2 obras. Observe que la primera obra puede ser asignada de 4 maneras (hay 4 residentes disponibles) y la segunda obra puede ser asignada de 3 maneras (ya se asignó 1 residente a la primera). Por lo tanto, el número total de asignaciones es: 4 x 3 = 12 maneras. Observe que 4x3 = 4x3x2x1 / 2x1 = 4! / 2!, usando la notación se tiene:

$${}_4P_2 = 4! / (4 - 2)! = 12.$$

- **COMBINACIONES:** El número de combinaciones de n objetos tomando m a la vez (para $m \leq n$) y denotado como: ${}_nC_m = n! / m! (n - m)!$

  Las 12 permutaciones del ejemplo anterior, denotando los residentes con las letras A, B, C y D y las obras con los dígitos 1 y 2, son: [A→ 1 y B → 2], [B → 1 y A → 2], [A→ 1 y C→ 2], [C → 1 y A → 2], [A →1 y D → 2], [D →1 y A → 2], [B→ 1 y C→ 2], [C→ 1 y B→ 2], [B → 1 y D → 2], [D → 1 y B → 2], [C → 1 y D → 2] y [D → 1 y C → 2].

  Observe que las permutaciones [A → 1 y B → 2] y [B → 1 y A → 2] involucran a la pareja de residentes A y B pero alternando la obra asignada a cada residente de la pareja, así, estas dos permutaciones constituyen una combinación. Esto es, en las combinaciones no importa el orden de los arreglos..

  Ejemplo 7: Suponga en relación con el ejemplo anterior que el gerente de la compañía desea saber el número total de parejas diferentes que pueden asignarse, sin importar a cuál de las obras se asigne cada residente de la pareja.

  Objetivo del ejemplo 7: Mostrar la aplicación de las combinaciones para determinar el número de resultados posibles de un experimento, además de la diferencia entre permutaciones y combinaciones.

Solución:

Esto implica que las dos primeras duplas de la lista de permutaciones [A→1 y B→ 2] y [B → 1 y A → 2] son equivalentes y por lo tanto las 2 duplas se transforman en la dupla única: [A y B], la cual, es una de las combinaciones de los cuatro residentes tomados de dos en dos a la vez.

Lo mismo puede decirse de las duplas 3 y 4, y así sucesivamente.

Por lo tanto, el total de parejas posibles será:

$$_4C_2 = 4! \,/\, 2!\,(4 - 2)! = 6 \text{ combinaciones.}$$

Estas parejas son: (A – B); (A – C); (A – D); (B – C); (B – D) y (C – D)

5. TALLER

1. La familia Uribe está compuesta por: Luis, Dora, René, Lina, Inés, Rita, Pepe, Roque, José y Guido. Luis, Dora, Inés, Pepe y Roque poseen carro y Luis, René, Rita y Dora poseen moto.

- Determine el conjunto universal (llámelo: U)
- Determine el conjunto asociado al suceso “tienen carro” (llámelo: C)
- Determine el conjunto asociado al suceso “tienen moto” (llámelo: M)
- Determine los conjuntos asociados a los siguientes sucesos:
    - No tienen carro (llámelo NC)
    - No tienen moto (llámelo NM)

- Determine los conjuntos asociados con los siguientes sucesos y, además, expréselos como operaciones entre conjuntos
    - Tienen carro y moto
    - No tienen ni carro ni moto
    - Tienen carro o tienen moto

- Determine los conjuntos asociados a los siguientes sucesos y, además, expréselos como operaciones entre conjuntos:
    - Tienen carro pero no tienen moto
    - Tienen moto pero no tienen carro
    - No tienen moto o no tienen carro

2. Cuantas placas de auto de cinco dígitos diferentes pueden hacerse con los dígitos del 0 al 9?

3. Si las placas tienen dos letras y cuatro dígitos, cuantas placas diferentes pueden hacerse con las vocales (ocupando las dos primeras posiciones) y los dígitos del 0 al 9?

4. Los dígitos del cero al 9 salieron de paseo, los múltiplos de tres sólo hablan chino y los demás sólo español. Una noche llegaron a un hotel que sólo tenía disponible una habitación con dos camas.

   - De cuántas maneras podrán ocupar las camas?
   - De cuántas maneras podrán ocupar la habitación?
   - De cuantas maneras podrán ocupar la habitación de manera que sus ocupantes puedan comunicarse durante la noche?

5. A una isla desierta llegan ocho náufragos, 4 mujeres y 4 hombres.

   - Si deciden construir ocho chozas, de cuántas maneras podrán ocuparlas?
   - Si deciden construir cuatro chozas dobles (para dos), de cuántas maneras podrán ocuparlas?
   - Si deciden construir dos chozas con capacidades de cinco y tres ocupantes, de cuántas maneras podrán ocuparlas?

6. Dos barcos pesqueros faenan en un lago. Estos se comunican usando un juego de cinco banderas (Roja, verde, azul, blanca y negra). Y un "código" de dos banderas, izando una bandera arriba y la otra abajo en el mástil.

   - Cuántos mensajes diferentes son posibles?

7. En relación con el problema anterior, determine el número posible de mensajes en las siguientes situaciones:

   - Con 5 banderas y el código de 3 banderas?
   - Con 6 banderas y el código de 2 banderas?
   - Con 6 banderas y el código de 3 banderas?

8. Dos suicidas deciden jugar a la ruleta rusa. En un revolver de seis tiros se pone una bala en el tambor y alternativamente accionan el gatillo girando el tambor previamente.

- Cuál es el espacio de muestras de este "experimento" al hacer dos rondas (cada uno acciona el gatillo 2 veces).

- Determine el espacio de muestras si no se gira el tambor antes de accionar el gatillo, al hacer dos rondas.

9. En un lote de 50 lápices hay 18 defectuosos.

- De cuántas maneras diferentes pueden extraerse cuatro lápices del lote?
- De cuántas maneras diferentes pueden extraerse dos lápices defectuosos?
- De cuántas maneras diferentes pueden extraerse dos lápices buenos?
- De cuántas maneras diferentes pueden extraerse dos lápices defectuosos y dos lápices buenos?

10. En el juego del Baloto se extraen al azar 6 balotas numeradas del 1 al 42. De cuántas maneras pueden extraerse las 6 balotas?

11. Una lotería juega con cinco ruedas, las tres primeras ruedas están numeradas del 0 al 9, la cuarta con los 12 signos zodiacales y la quinta con las cinco vocales. Cuántos billetes diferentes deberán emitirse para su sorteo?

12. Al torneo de tenis de las Américas, se inscribieron 16 jugadores, de los cuales 2 son colombianos. Los jugadores enfrentados en la primera ronda se eligen al azar y pasan a la segunda ronda los ganadores, quedando eliminados los perdedores.

- De cuántas maneras diferentes pueden quedar los jugadores que clasifiquen a la segunda ronda?
- De cuántas maneras diferentes los 2 colombianos, pueden quedar entre los clasificados a la segunda ronda?

6. BIBLIOGRAFIA

- Meyer, P. *Probabilidad y Aplicaciones Estadísticas.* Fondo Educativo Interamericano. México 1981.
- Ash, R. *Basic Probability Theory.* John Wiley & Sons. New York. 1976.
- Walpole, R. *Introduction to Statistics*. Mc. Millan Company. New York. 1969
- Shao Stephen. *Estadística para Economistas y Administradores de Empresas*. Herrero Hermanos. México. 1979
- Gómez L. Hernán. *Estadística Experimental*. Universidad Nacional de Colombia. Bogotá. 1997

MODULO 2: PROBABILIDAD

OBJETIVOS DEL MODULO:

- Comprender el concepto de probabilidad
- Comprender el concepto de probabilidad condicional
- Comprender la Regla de Bayes
- Relacionar los conceptos de probabilidad con actividades productivas y empresariales.

CONTENIDO:

PRERREQUISITOS:

- Modulo 1. Espacios de Muestras y Sucesos

1. FRECUENCIA RELATIVA:

Si un experimento se repite k veces y el suceso A asociado al experimento ocurre p veces en las k repeticiones, la frecuencia relativa de ocurrencia del suceso A, se define como: $f_A = p / k$ y presenta las siguientes características:

- $0 \leq f_A \leq 1$
- $f_A = 1$ Si y sólo si el suceso A ocurre en las k repeticiones.
- $f_A = 0$ Si y sólo si el suceso A nunca ocurre en las k repeticiones.

Ejemplo 1: Considere el siguiente experimento, en la bifurcación de una transitada carretera se registra la dirección que toma cada vehículo (hacia A o B). Un día cualquiera llegaron a la bifurcación 670 vehículos de los cuales 235 tomaron la ruta hacia A (suceso A). ¿Cuál es la frecuencia relativa del suceso A?

Objetivo del ejemplo: Mostrar la aplicación de la frecuencia relativa de un suceso asociado a la repetición de un experimento.

Solución:

Experimento: Registrar la ruta (A o B) que toma un vehículo al llegar a la bifurcación.

Repeticiones del experimento: El número de vehículos que llegan a la bifurcación, en este caso k = 670

Suceso A: Número de vehículos que toman la ruta A, en este caso p = 235.

Frecuencia relativa del suceso A: f_A = 235 / 670 = 0.3507

Observe que si el número de repeticiones del experimento (bajo las mismas condiciones) se incrementa, digamos al número de vehículos que llegan a la bifurcación durante un año, la frecuencia relativa de la ocurrencia del suceso A mostrará un patrón de "regularidad".

Suponga en relación con el ejemplo 1 que en un año cualquiera de 240358 vehículos, 86012 tomaron la ruta hacia A, caso en el cual: f_A = 86012 / 240358 = 0.3578.

El patrón de regularidad de la frecuencia relativa del suceso A para un número de repeticiones que tienda a infinito puede asociarse a la probabilidad de ocurrencia del suceso.

2. PROBABILIDAD

La probabilidad de ocurrencia del suceso A asociado a un experimento cuyo espacio de muestras consta de n resultados diferentes, y denotada como: P(A), se define como la relación del número posible de resultados del suceso A (n_A) y el número total de resultados del experimento (n), es decir, P(A) = n_A / n.

La validez de esta formula implica la suposición de que la ocurrencia de cada uno de los n resultados del experimento están sometidos a iguales condiciones.

Ejemplo 2: Considere el experimento de lanzar un dado y registrar el número de puntos en que cae. ¿Cuál es la probabilidad de caiga en uno?

Objetivos del ejemplo: Mostrar la aplicación del concepto de probabilidad y su relación con los conceptos de espacio de muestras y suceso.

Solución:

Espacio de Muestras: S = [1, 2, 3, 4, 5, 6} Número de resultados n = 6

Suceso A: A = {1} Número de elementos del suceso $n_A = 1$

$$P(A) = 1 / 6$$

Observe que la validez de este resultado está condicionado a que el dado sea equilibrado, es decir que, cada uno de los posibles resultados puedan ocurrir (iguales condiciones)

3. PROPIEDADES DE LA PROBABILIDAD:

La P(A) tiene las siguientes propiedades:

- $0 \leq P(A) \leq 1$
  P(A) = 1, si y sólo si el suceso A siempre ocurre, es decir A = S.
  P(A) = 0, si y sólo si el suceso nunca ocurre.

- P(S) = 1
  La probabilidad de ocurrencia de todos los elementos del espacio muestral o sea, de todos los posibles resultados del experimento es uno.

- Si A y B son sucesos que se excluyen mutuamente (no pueden ocurrir simultáneamente), la $P(A \cup B) = P(A) + P(B)$.

- $P(\emptyset) = 0$

- Si los sucesos A y B son independientes, la probabilidad de ocurrencia de A y B es el producto de sus probabilidades: $P(A \cap B) = P(A) \times P(B)$

- $P(A) = 1 - P(A^C)$

- Si A y B son dos sucesos cualesquiera la $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Ejemplo 3: Considere el experimento de lanzar un dado equilibrado. Calcule la probabilidad de ocurrencia de los siguientes sucesos:

- Que caiga 5 o 6
- Que caiga en un número par de puntos
- Que caiga en 7
- Que no caiga en 3.
- Que caiga en 2 o en un número par de puntos.

Objetivo del ejemplo: Mostrar la aplicación de las propiedades de la probabilidad.

Solución:

- $P(5 \text{ o } 6) = P(5 \cup 6) = P(5) + P(6) = 1 / 6 + 1 / 6 = 2 / 6 = 1/3$
- $P(\text{número par de puntos}) = P(2 \cup 4 \cup 6) = 1/6 + 1/6 + 1/6 = 3/6 = 1/2$
- $P(7) = 0 / 6 = 0$ (no puede ocurrir)
- $P(\text{no caiga en } 3) = 1 – P(3) = 1 – 1/6 = 5/6$.
- $P(2 \text{ o Par}) = P\{2 \cup \text{Par}\} = P(2) + P(\text{Par}) – P(2 \cap \text{Par})$, se tiene: $P(2) = 1/6$ $P(\text{Par}) = 1/2$, y $(2 \cap \text{Par}) = (2)$, por lo tanto, $P(2 \cap \text{Par}) = 1/6$ y la probabilidad pedida, es: $1/6 + 1/2 - 1/6 = 1/2$

Ejemplo 4: Considere el experimento de lanzar un dado y una moneda. Calcule la probabilidad de ocurrencia de los siguientes sucesos:

- Que el dado caiga en 1 y la moneda en cara
- Que el dado caiga en 1 o la moneda caiga en cara

Objetivo del ejemplo: Mostrar la aplicación de las propiedades de la probabilidad.

Solución:

- Dado que los resultados obtenidos en el dado y la moneda son independientes, la probabilidad pedida, es: $P(1) \times P(\text{Cara}) = 1/6 \times 1/2 = 1/12$

- $P(1 \cup \text{Cara}) = P(1) + P(\text{Cara}) – P(1 \cap \text{Cara}) = 1/6 + 1/2 – 1/12 = 7/12$

4. PROBABILIDAD CONDICIONAL

En situaciones, como la del ejemplo 3A, la ocurrencia del suceso, el dado cae en 1, no afecta la ocurrencia de cómo caiga la moneda, en otras palabras, los sucesos son independientes. Sin embargo, en ocasiones, la ocurrencia de un suceso está condicionada a la ocurrencia de un suceso previo.

Como ilustración considere la dependencia de la ocurrencia de inundaciones en verano y la cantidad de nieve caída en la época invernal precedente.

A partir de los ejemplos 4 y 5, se ilustrará la diferencia entre las probabilidades de ocurrencia de sucesos independientes y sucesos dependientes, así como, la definición de probabilidad condicional.

Ejemplo 5: Considere un lote de 10 artículos de los cuales 3 son defectuosos. Si se eligen 2 artículos del lote, uno a continuación del otro con sustitución, es decir se mira el primer artículo y se regresa al lote.

- Cuál es la probabilidad de que el primer articulo sea defectuoso?
- Cuál es la probabilidad de que el segundo articulo sea defectuoso?

Objetivo del ejemplo: Ilustrar el concepto de sucesos independientes

Solución:

Se definen los sucesos A y B, así:

A = { El primer artículo es defectuoso }
B = { El segundo artículo es defectuoso }

Las probabilidades de los sucesos son: P(A) = P(B) = 3/10

Observe que las dos repeticiones del experimento se hacen bajo las mismas condiciones, es decir el espacio de muestras del experimento no cambia en las dos repeticiones, y por lo tanto, el resultado de la segunda repetición es independiente de lo ocurrido en la primera.

Ejemplo 6: En relación con el problema 5, calcule las probabilidades de los sucesos A y B, pero sin sustitución, es decir, si el articulo observado en la primera extracción no se regresa al lote.

Objetivo del ejemplo: Mostrar las diferencias entre sustituir o no el artículo de la primera extracción, en relación con sus espacios de muestras.

Solución:

En el caso de no sustituir el artículo de la primera extracción al lote se tiene:

$$P(A) = 3/10.$$

Para calcular la probabilidad del suceso B es necesario saber cuál fue el resultado de la primera extracción, dado que este resultado determina la nueva composición del espacio de muestra para la segunda extracción, así:

- Si el primer articulo fue defectuoso: El lote tiene ahora 9 artículos de los cuales 2 son defectuosos y por lo tanto $P(B) = 2/9$.
- Pero si el primer articulo fue bueno: El lote tiene ahora 9 artículos de los cuales 3 son defectuosos y por lo tanto $P(B) = 3/9$.

Observe que a diferencia del resultado obtenido en el ejemplo 4, la probabilidad de ocurrencia del suceso B, está condicionada a si ocurrió o no el suceso A en la primera extracción.

De la situación anterior se define la probabilidad condicional del suceso B dado que el suceso A ha ocurrido, y denotada como: $P(B / A) = P(A \cap B) / P(A)$. Análogamente la probabilidad de ocurrencia del suceso B, dado que el suceso A no ocurrió (A^C), es: $P(B / A^C) = P(A^C \cap B) / P(A^C)$.

Ejemplo 7: La siguiente tabla muestra las características de un grupo de 40 personas:

	Hombres (H)	Mujeres (M)	TOTALES
Casados (C)	11	8	19
Solteros (S)	7	14	21
TOTALES	18	22	40

Tabla 2.1. - Datos ejemplo 6.

Si se elige 1 persona del grupo al azar y resulta ser una mujer, cuál es la probabilidad de que sea casada?

Objetivo de los ejemplos 7 y 8: Mostrar la aplicación del concepto de probabilidad condicional.

Solución:

La probabilidad pedida es: $P(C / M) = P(C \cap M) / P(M)$
De la tabla pueden calcularse:

$P(C \cap M) = 8 / 40$ y $P(M) = 22 / 40$
Por lo tanto, $P(C / M) = (8 / 40) / (22 / 40) = 0.36$

Observe que esta probabilidad podría haberse calculado considerando el "estado del lote" , luego de conocer el sexo de la persona elegida, es decir, dado que ésta fue una mujer, el nuevo espacio de muestras es 22 de las cuales 8 son casadas, por lo tanto, $P(C / M) = 8 / 22 = 0.36$.

Ejemplo 8: En un lote de pernos de 100 unidades hay 6 defectuosos (D), se extraen 2 pernos al azar (sin sustitución). Cuál es la probabilidad de que ambos artículos sean defectuosos?

Sean los sucesos:

A = { El primer artículo es defectuoso }
B = { El segundo artículo es defectuoso }

Se pide $P(A \cap B)$, de la definición de probabilidad condicional se tiene:
$P(A \cap B) = P(B / A) \times P(A)$

Pueden calcularse: $P(A) = 6/100$ y $P(B / A) = 5/99$, por lo tanto la probabilidad pedida es: $P(A \cap B) = (5/99) (6/100) = 0.303\%$

5. **REGLA DE BAYES:** Si $B_1, B_2, B_3, ... B_k$, es un conjunto de sucesos mutuamente excluyentes y su unión es igual al conjunto universal (espacio de muestras). Y el suceso A es un subconjunto del conjunto universal o espacio de muestras (Figura 2.1), entonces:

$$P(B_K / A) = \frac{P(B_K \cap A)}{P(B_1 \cap A) + P(B_2 \cap A) + \ldots + P(B_K \cap A)}$$

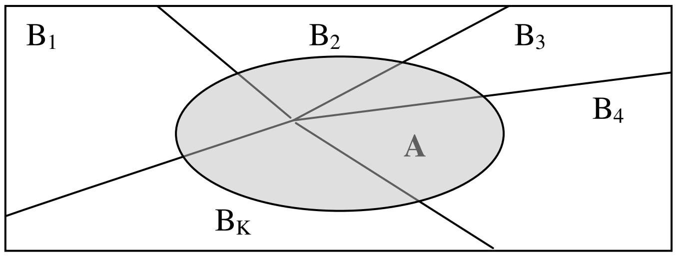


Figura 2.1. - Diagrama de Venn, sucesos B_K y A.

Ejemplo 9: El número de estudiantes pertenecientes a las facultades de Minas, Agronomía y Ciencias en tres grupos de estadística se muestra en la siguiente tabla:

	Grupo 1	Grupo 2	Grupo 3
MINAS (M)	18	14	12
AGRONOMIA (A)	7	20	10
CIENCIAS (C)	20	12	16

Tabla 2.2. - Datos del ejemplo 8

Si se selecciona al azar uno de los tres grupos y de este se elige un estudiante al azar resultando que pertenece a la facultad Minas, ¿Cuál es la probabilidad de que pertenezca al grupo 2?

Objetivo del ejemplo: Mostrar la aplicación de la regla de Bayes.

Solución:

Se pide $$P(G2 / M) = \frac{P(G2 \cap M)}{P(G1 \cap M) + P(G2 \cap M) + P(G3 \cap M)}$$

Se conocen:

$P(G1) = P(G2) = P(G3) = 1 / 3$
$P(M / G1) = 18 / 45$; $P(M / G2) = 14 / 46$; $P(M / G3) = 12 / 38$;

Por lo tanto pueden hallarse:

$P(G1 \cap M) = P(M / G1) \times P(G1) = (18/45)x(1/3) = 0.1333$
$P(G2 \cap M) = P(M / G2) \times P(G2) = (14/46)x(1/3) = 0.1014$
$P(G3 \cap M) = P(M / G3) \times P(G3) = (12/38)x(1/3) = 0.1053$

Entonces $P(G2 / M) = \dfrac{0.1014}{0.1333 + 0.1014 + 0.1053} = 29.82\%$

Adicionalmente, $P(G1/ M) = 39.20\%$ y $P(G3/ M) = 30.98\%$

Observe que: $P(G1/ M) + P(G2/ M) + P(G3/ M) = 100\%$

6. TALLER

1. En una ciudad se tienen registros históricos del régimen de lluvias. Si en los últimos 500 días hubo 73 días lluviosos. ¿Cuál es la probabilidad de que mañana llueva?

2. En cada una de tres cajas hay 100 pernos de los cuales 18 son defectuosos. Si se elige al azar 1 perno de cada caja.

- ¿ Cuál es la probabilidad de que los tres pernos sean defectuosos?
- ¿Cuál es la probabilidad que el de la primera caja sea defectuoso y los otros dos sean buenos?
- ¿Cuál es la probabilidad de que dos de los tres pernos sean defectuosos?

3. En una caja hay 5 balotas marcadas con los números del uno al cinco. ¿Cuál es la probabilidad de que al extraer, al mismo tiempo, tres balotas resulten marcadas con los números uno, dos y tres?

4. Una empresa compra un tipo de perno a tres proveedores (P1, P2 y P3). El proveedor 1 sirve los pernos en lotes de 1000 unidades y con una fracción defectuosa del 15%. El proveedor 2 en lotes de 500 y 12% defectuosos. El proveedor 3 en lotes de 800 y 8% defectuosos.

Un día llegaron tres lotes, uno de cada proveedor y se vaciaron a un contenedor, el gerente llegó en ese momento y tomó al azar un perno del contenedor , lo examino y viéndolo defectuoso, preguntó: ¿De cuál proveedor será este perno?

¿Qué podría contestársele al gerente?

5. En una caja hay dos bolas blancas, en una segunda caja hay dos bolas negras y en la tercera hay una bola blanca y una negra. Se elige una de las cajas al azar y se toma una bola resultando ser blanca. ¿Cuál es la probabilidad de que la otra bola de esa misma caja también sea blanca?

6. Dos suicidas deciden jugar a la ruleta rusa. En un revolver de seis tiros se pone una bala en el tambor y alternativamente accionan el gatillo girando el tambor previamente.

- Calcule la probabilidad de que ambos terminen ilesos si juegan dos rondas (cada uno acciona el gatillo dos veces).
- Calcule la probabilidad de que ambos terminen ilesos si juegan dos rondas pero no se gira el tambor.

7. El “sardino” Sierra, sabe que si la noche anterior llueve la probabilidad de que la pesca sea buena es del 90%. Además se sabe que el 40% de las noches llueve.

- Cuál es la probabilidad de que la víspera llueva y la pesca sea buena?

8. En otro lago se sabe que cuándo llueve la víspera la probabilidad de que la pesca sea buena es del 75%. Además que la probabilidad de que la víspera haya llovido dado que la pesca fue buena es del 65%. Y finalmente que la probabilidad de que haya llovido la víspera y la pesca sea buena es del 55%.

- Cuál es el régimen de lluvias en este lago?
- En los próximos cien días cuántos días se esperarían que fueran de buena pesca?

9. La Caverna de los Baules es un programa de concurso. El anfitrión dirigiéndose a un concursante le dice: "En la caverna hay tres baúles usted debe elegir uno de ellos. Cada baúl tiene un cheque, uno de ellos es de $ 100.000, otro de $ 600.000 y el otro de $ 900.000. Haga su elección... .

Felicitaciones! acaba de ganar $ 600.000. Sin embargo le propongo el siguiente negocio: Si quiere puede llevarse sus $ 600.000 o hacer una segunda elección entre los dos baúles que quedan en la caverna, pero ahora a cada uno de los baúles le vamos a remplazar el cheque por dos cheques uno por el doble y otro por la mitad del valor que tenían originalmente y elegir del baúl al azar uno de los cheques "

- Que debe hacer el concursante?

10.-Dolores una mañana al despedirse de sus hijos les encaró comprar algunas pastillas para el dolor de cabeza puesto que las había acabado. De regreso a casa Carlos le entregó una bolsa con 20 caspirinas, Saúl una bolsa con 50 saspirinas y Raúl una bolsa con 45 raspirnas. Dolores vacío las tres bolsas de pastillas a un cajón de su mesa de noche.

Durante la cena Carlos le pregunto cual de los tres tipos de pastillas prefería, a lo que Dolores contestó: "La caspirina me calma el dolor el 75% de las veces, la Saspirina el 45% mientras que la raspirina lo hace el 65% de las veces".

Al día siguiente Dolores le contó a sus hijos que un fuerte dolor la había despertado pero que tomó del cajón la primera pastilla que se le atravesó calmándole el dolor de cabeza.

Carlos intervino diciendo que con seguridad había tomado una caspirina puesto que es la más efectiva. Saúl, por su parte, dijo que el creía que había tomado una saspirina puesto que éstas eran las mas numerosas en el cajón. Finalmente, Raúl intervino lamentando no estar de acuerdo con sus hermanos y manifestando tener razones para creer que Dolores había tomado raspirina.

- A que razones se refería Raúl para justificar su afirmación?

11. En el juego de Loto se extraen seis balotas al azar numeradas del 1 al 42. Quien acierte los seis números en cualquier orden es el ganador del premio

- Cuál es la probabilidad de acertar los seis números?
- Cuál es la probabilidad de que los seis números extraídos sean pares?

12. Ana, Beto, Ciro, Dora y Emma se forman en fila.

- Cuál es la probabilidad de que se ubiquen en orden alfabético ascendente?
- Cuál es la probabilidad de que se ubiquen en orden alfabético descendente?
- Cuál es la probabilidad de que se ubiquen en orden alfabético ascendente o descendente?

13. Ocho tenistas, de los cuales, dos son colombianos, se enfrentarán en cuartos de final, por parejas escogidas al azar, quedando eliminados los perdedores, bajo el mismo sistema los clasificados disputarán el derecho a jugar la final del torneo. ¿Cuál es la probabilidad de que el partido final lo disputen los dos colombianos?

7. BIBLIOGRAFIA

- Meyer, P. *Probabilidad y Aplicaciones Estadísticas.* Fondo Educativo Interamericano. México 1981.
- Ash, R. *Basic Probability Theory.* John Wiley & Sons. New York. 1976.
- Walpole, R. *Introduction to Statistics*. Mc. Millan Company. New York. 1969
- Shao Stephen. *Estadística para Economistas y Administradores de Empresas*. Herrero Hermanos. México. 1979
- Gómez L. Hernán. *Estadística Experimental*. Universidad Nacional de Colombia. Bogotá. 1997

MODULO 3 VARIABLES ALEATORIAS Y DISTRIBUCIONES DE PROBABILIDAD

OBJETIVOS DEL MODULO:

- Comprender el concepto de Variable Aleatoria
- Identificar los tipos de Variables Aleatorias.
- Comprender el concepto de Distribución de Probabilidad, sus características y propiedades.
- Comprender los conceptos de Valor Esperado y Varianza de una variable aleatoria conocida su distribución de probabilidad
- Aprender la aplicación práctica de los anteriores conceptos en relación con actividades productivas y empresariales

CONTENIDO:

PRERREQUISITOS:

- Estadística descriptiva
- Módulos 1 y 2

1. VARIABLES ALEATORIAS

La variable aleatoria X, se define como, la función X(s) que asigna un número real a cada uno de los elementos del espacio de muestras S, esto es, a cada uno de los posibles resultados de un experimento. Observe que el concepto de variable aleatoria está ligado a la totalidad de los sucesos asociados con el experimento.

Como ilustración, considere el siguiente espacio de muestras, S = {Cali, Tolú, Bello, Medellín, Bogotá, Quibdo}. Si se define la variable aleatoria X, como el número de letras de cada una de las ciudades de S, se tiene:

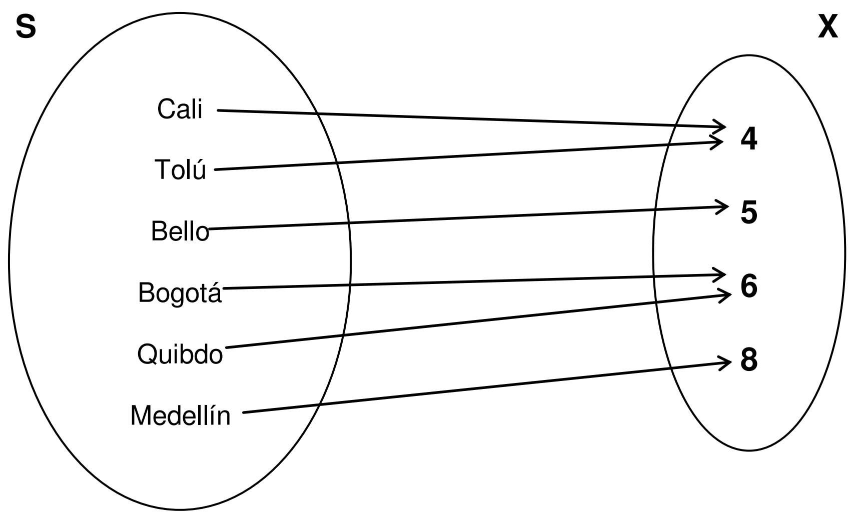


Figura 3.1 Variable aleatoria X(s)

Observe que para cada uno de los elementos s del espacio de muestras S, se asigna el valor x del conjunto X, como el número de letras de cada elemento s.

Ejemplo 1: Considere un lote de 10 artículos de los cuales 3 son defectuosos y 7 son buenos. Se extraen 2 artículos en sucesión sin sustitución determinando el número de artículos defectuosos extraídos.

Objetivo del ejemplo: Mostrar la relación entre el espacio de muestras del experimento y la variable aleatoria definida.

Solución:

El espacio de muestras de este experimento, denotando b = artículo bueno y d = artículo defectuoso, es: S = $\{$ bb, bd, db, dd $\}$.

La variable aleatoria X se definió como el numero de artículos defectuosos obtenidos, y por lo tanto, se tiene:

X(bb) = 0; X(bd) = 1; X(db) = 1; X(dd) = 2

Así, X = { 0, 1, 2}

VARIABLES ALEATORIAS DISCRETAS: Una variable aleatoria X, es discreta, si sus elementos x, son valores discretos. La variable aleatoria X, definida en el ejemplo 1 es discreta.

VARIABLES ALEATORIAS CONTINUAS: En el caso de que la variable aleatoria represente una magnitud en una escala continua, tales como: longitud, peso, área, volumen, densidad, etc. Se dice que la variable aleatoria es continua.

Considere la variable aleatoria X, como el diámetro de las canicas producidas por una máquina, así: X = { x / x > 0}. Observe que los diámetros son números reales mayores que cero y por lo tanto, habrá infinitos valores de su diámetro, y en consecuencia X es una variable aleatoria continua.

2. DISTRIBUCIÓN DE PROBABILIDAD

Si a cada uno de los posibles valores x_i de la variable aleatoria X se le asocia su probabilidad de ocurrencia $P(X = x_i) = p(x_i)$, las duplas [x_i , $p(x_i)$], para todo i, se obtiene la distribución de probabilidad de la variable aleatoria X. Se dice que la distribución de probabilidad es discreta o continua, si la variable aleatoria es discreta o continua. En este modulo sólo se tratarán las variables aleatorias discretas; las continuas se examinarán en el módulo 5.

Como se vio en el módulo 2, los valores $p(x_i)$, deben cumplir las siguientes condiciones:

- $0 \leq p(x_i) \leq 1$ para todo i
- La suma de las probabilidades de ocurrencia de la totalidad de los resultados del espacio de muestras es 1.

Para el caso de distribuciones de probabilidad discretas, se tiene:

$$\sum_{i=1}^{\infty} p(x_i) = 1$$

Ejemplo 2: Se giran dos ruletas numeradas del 1 al 4 y se anotan los valores obtenidos en cada una. Si se define la variable aleatoria, X(s) = x, como la suma de los valores obtenidos en ambas ruletas, determine la distribución de probabilidad de X.

Objetivo del ejemplo 2: Mostrar como determinar la distribución de probabilidad de la variable aleatoria y validar el cumplimiento de las condiciones citadas.

Solución: El espacio de muestras de este experimento tiene 16 elementos, así:

S = {1-1, 1-2, 1-3, 1-4, 2-1, 2-2, 2-3, 2-4, 3-1, 3-2, 3-3, 3-4, 4-1, 4-2, 4-3, 4-4}

Los valores de la variable aleatoria, son:

X(1-1) = 2
X(1-2) = X(2-1) = 3
X(1-3) = X(2-2) = X(3-1) = 4
X(1-4) = X(2-3) = X(3-2) = X(1-4) = 5
X(2-4) = X(3-3) = X(4-2) = 6
X(3-4) = X(4-3) = 7
X(4-4) = 8

Las probabilidades asociadas a cada uno de los valores de la variable aleatoria X, se muestran en la siguiente tabla:

X(s) = x	2	3	4	5	6	7	8
P(X = x) = p(x)	1/16	2/16	3/16	4/16	3/16	2/16	1/16

Tabla 3.1. - Distribución de probabilidad de la variable aleatoria X

Dado que la variable aleatoria X es discreta su distribución de probabilidad también lo es.

Además, la suma de las probabilidades de la totalidad del espacio de muestras es 1, así:

$$\sum_{i=1}^{\infty} p(X_i) = 1/16 + 2/16 + 3/16 + 4/16 + 3/16 + 2/16 + 1/16 = 1$$

3. **VALOR ESPERADO:** El valor esperado, esperanza o media, E(X), de una variable aleatoria X discreta con la siguiente distribución de probabilidad :

X(s) = x	x_1	x_2	x_3	....	x_n
P(X = x) = p(x)	$p(x_1)$	$p(x_2)$	$p(x_3)$	....	$p(x_n)$

Tabla 3.2. - Distribución de probabilidad de la variable aleatoria X

Se define como: $E(X) = \sum_{i=1}^{n} x_i \, p(x_i)$

Ejemplo 3: Cuál es el valor esperado de la suma de los valores obtenidos en las ruletas del ejemplo 2?

Objetivo del ejemplo 3: Ilustrar el calculo e interpretación del valor esperado de la variable aleatoria discreta X

Solución:

$E(X) = 2(1/16) + 3(2/16) + 4(3/16) + 5(4/16) + 6(3/16) + 7(2/16) + 8(1/16) = 5$

Es decir, el valor promedio de la variable aleatoria es 5.

4. **VARIANZA:** La varianza, σ^2 , de una variable aleatoria discreta X, está dada por:

$$\sigma^2 = E(X - \mu)^2 = E(X^2) - \mu^2 \quad \text{donde: } \mu = E(X)$$

Ejemplo 4: Cuál es la varianza de la suma de los valores obtenidos en las ruletas del ejemplo 2?

Objetivo del ejemplo 4: Ilustrar el calculo e interpretación de la varianza de la variable aleatoria discreta X

Solución:

$E(X^2) = 4(1/16)+9(2/16)+16(3/16)+25(4/16)+36(3/16)+49(2/16)+64(1/16) = 27.3$

Por lo tanto, $\sigma^2 = 27.3 - 25 = 2.3$

La desviación estándar σ, es la raíz cuadrada de la varianza.

La desviación estándar indica el grado de variabilidad o la dispersión de los valores de la variable aleatoria en relación con la media.

En este caso, la desviación estándar es: $\sigma = 1.516$

Ejemplo 5: Un lote de 30 condensadores contiene 18 unidades defectuosas. Considere el experimento de extraer al azar 5 condensadores y registrar el número de unidades defectuosas obtenidas en la muestra.

Si se define la variable aleatoria X, como el número de defectuosos obtenidos.

- Qué valores puede tomar X?
- Determine la distribución de probabilidad de la variable aleatoria X
- Calcule la media y la varianza de la variable aleatoria X

Objetivo del ejemplo 5: Ilustrar la aplicación de los conceptos: de variable aleatoria, probabilidad y distribución de probabilidad.

Solución:

- X = (0, 1, 2, 3, 4, 5)

- Para determinar la distribución de probabilidad de la variable aleatoria X, es necesario el calculo de las probabilidades de ocurrencia de cada uno de los valores de X, así:

  Número de elementos del espacio muestral: ${}_{30}C_5 = 142506$

  Número de maneras de obtener:

Cero defectuosos:	${}_{18}C_0$	x	${}_{12}C_5$	=	1	x	792	=	792
Un defectuoso:	${}_{18}C_1$	x	${}_{12}C_4$	=	18	x	495	=	8910
Dos defectuosos:	${}_{18}C_2$	x	${}_{12}C_3$	=	153	x	220	=	33660
Tres defectuosos:	${}_{18}C_3$	x	${}_{12}C_2$	=	816	x	66	=	53856
Cuatro defectuosos:	${}_{18}C_4$	x	${}_{12}C_1$	=	3060	x	12	=	37720
Cinco defectuosos:	${}_{18}C_5$	x	${}_{12}C_0$	=	8568	x	1	=	8568

Probabilidades:

P(0 defectuosos) = 792 / 142506 = 0.56%
P(1 defectuoso) = 8910 / 142506 = 6.25%
P(2 defectuosos) = 33660 / 142506 = 23.62%
P(3 defectuosos) = 53856 / 142506 = 37.79%
P(4 defectuosos) = 37720 / 142506 = 25.77%
P(5 defectuosos) = 8568 / 142506 = 6.01%

Suma 100%

Distribución de probabilidad:

X = x	0	1	2	3	4	5
P(X = x)	0.56%	6.25%	23.62%	37.79%	25.67%	6.01%

- E(X) =1x6.25% + 2x23.62% + 3x37.79% + 4x25.67% + 5x6.01% = 3
- $E(X^2)$=1x6.25%+4x23.62%+9x37.79%+16x25.67%+25x6.01% = 10.0345
  Var(X) = 10.0345 – 9 = 1.0345

5. TALLER

1. Para cada uno de los siguientes experimentos determine los valores de la variable aleatoria correspondiente:

- Experimento: Lanzar un dado.
  Variable aleatoria: Puntos obtenidos.
- Experimento: Lanzar dos dados.
  Variable aleatoria: Suma de puntos obtenidos.
- Experimento: Lanzar dos dados.
  Variable aleatoria: Producto de los puntos obtenidos.
- Experimento: Lanzar dos dados.
  Variable aleatoria: Valor absoluto de la diferencia de puntos obtenidos.

- Experimento: Lanzar tres monedas
  Variable aleatoria: Número de caras obtenidas

- Experimento: Lanzar una moneda y un dado
  Variable aleatoria: Suma del número de caras y puntos obtenidos en la moneda y el dado

2. Determine la distribución de probabilidad de cada una de las variables aleatorias especificadas en el problema 1

3. El espacio de muestras de un experimento es: S = {-3, -2, -1, 1, 2, 4, 5, 9}, determine los valores de las siguientes variables aleatorias, así como las correspondientes distribuciones de probabilidad.

- Valor obtenido

- El cuadrado del valor obtenido

- Valor absoluto del resultado obtenido

4. La siguiente tabla muestra las probabilidades de ocurrencia de la variable aleatoria discreta X.

Valor de X	0	7	8	11	12	14
Probabilidad	1/ 10	1/ 10	3/ 10	2/ 10	1/ 10	2/ 10

- Demuestre que esta tabla representa la distribución de probabilidad de la variable aleatoria X

- Calcule la media y la varianza de la variable aleatoria X

5. Una caja contiene 5 pernos de los cuales 2 son defectuosos y 3 son buenos, y se extraen sucesivamente pernos al azar sin sustitución. Se define la variable aleatoria X, como el número de pernos que es necesario extraer hasta hallar un perno defectuoso.

- Determine los valores que puede tomar la variable aleatoria
- Establezca la distribución de probabilidad de la variable aleatoria
- Calcule el valor esperado y la varianza de la variable aleatoria

6. Cuál es el valor esperado de una rifa de cuatro cifras cuyo premio es 1 millón de pesos al comprar una boleta que cuesta $ 500.

7. En un lote de semillas, se sabe que la probabilidad de que una semilla germine bajo ciertas condiciones es del 70%. Si se extraen 3 semillas al azar y se define la variable aleatoria X, como el número de semillas que germinan. Determine:

- La distribución de probabilidad de la variable aleatoria X
- La media de la variable aleatoria X
- La desviación estándar de la variable aleatoria X

8. La variable aleatoria X puede tomar los valores: 2, 3 y 6. Si la probabilidad de ocurrencia de la variable aleatoria es: $P(X = x_i) = 1 / x_i$

- Determine las probabilidades de ocurrencia de X
- Demuestre que $\{ x_i \, ; P(x_i) \}$ para $i = 1, 2, 3$, es una distribución de probabilidad.
- Calcule el valor esperado y la varianza de la variable aleatoria X

9. En un contenedor hay 15 guantes de boxeo, de los cuales 5 son derechos y 10 son izquierdos. Considere la variable aleatoria X, como el número de pares derecho – izquierdo obtenidos al elegir al azar 4 unidades del contenedor. Determine la distribución de probabilidad de la variable aleatoria X.

6. BIBLIOGRAFÍA

- Meyer, P. *Probabilidad y Aplicaciones Estadísticas.* Fondo Educativo Interamericano. México 1981.
- Ash, R. *Basic Probability Theory.* John Wiley & Sons. New York. 1976.
- Walpole, R. *Introduction to Statistics.* Macmillan Company. New York 1969.
- Mason, R. y D. Lind. *Estadística para Administración y Economía.* Alfaomega. México 1992

MODULO 4: DISTRIBUCIONES DISCRETAS DE PROBABILIDAD

OBJETIVOS DEL MÓDULO:

- Conocer las características y usos de las distribuciones de probabilidad discretas más usadas en el Control Estadístico de Calidad (Binomial, Hipergeométrica y Poisson)

CONTENIDO:

PRERREQUISITOS:

- Módulos 1, 2 y 3

1. LA DISTRIBUCIÓN BINOMIAL

Una de las distribuciones de probabilidad más usadas en el control estadístico de calidad es la distribución binomial.

Se define a un experimento como binomial, si cumple las siguientes condiciones:

- El experimento sólo tiene dos resultados posibles mutuamente excluyentes; sucesos A y A^C
- La probabilidad de ocurrencia del suceso A es p; $P(A) = p$, y por consiguiente, $P(A^C) = 1 – p$
- Si en n repeticiones independientes del experimento permanece constante la probabilidad de ocurrencia del suceso A

Si se define la variable aleatoria X, como el número de veces que ocurre el suceso A en las n repeticiones de un experimento binomial, con P(A) = p, se dice que la variable X es una variable aleatoria binomial con parámetros n y p. (denotada como b(n ; p))

La probabilidad de que ocurra k veces el suceso A en n repeticiones, está dada por:

$$P(X = k) = {}_{n}C_{k}\ p^{k}\ (1 - p)^{n-k} \qquad k = 0, 1, 2, \ldots\ldots, n$$

Ejemplo 1: Un lote de 10 pernos contiene 4 defectuosos. Considere el experimento de elegir un perno al azar del lote y luego de observar si es o no defectuoso se regresa al lote. Cuál es la probabilidad de que en tres repeticiones de este experimento se obtengan 0 defectuosos?

Objetivo del Ejemplo: Mostrar el cumplimiento de las condiciones de un experimento binomial y la aplicación de la distribución binomial.

Solución:

Este experimento tiene dos resultados posibles mutuamente excluyentes, esto es, el perno o es defectuoso o es no defectuoso (bueno) y la probabilidad de que sea defectuoso es 4/10 en cada una de las tres repeticiones. Por lo tanto, el experimento es binomial y la variable aleatoria X, definida como el número de pernos defectuosos obtenidos en las tres repeticiones, se distribuye binomial con parámetros n = 3 y p = 0.4 (X ~ b(3; 0.4))

En consecuencia, la probabilidad de obtener 0 defectuosos, es:

$$P(X = 0) = {}_{3}C_{0}\ p^{0}\ (1 - p)^{3-0} = 1 \times 1 \times (1 - 0.4)^{3} = 0.216 = 21.6\%$$

De manera análoga, las probabilidades de obtener 1, 2 y 3 defectuosos, están dadas, por:

$$P(X = 1) = {}_{3}C_{1}\ p^{1}\ (1 - p)^{3-1} = 3 \times 0.4^{1} \times (1 - 0.4)^{2} = 0.432 = 43.2\%$$
$$P(X = 2) = {}_{3}C_{2}\ p^{2}\ (1 - p)^{3-2} = 3 \times 0.4^{2} \times (1 - 0.4)^{1} = 0.288 = 28.8\%$$
$$P(X = 3) = {}_{3}C_{3}\ p^{3}\ (1 - p)^{3-3} = 1 \times 0.4^{3} \times (1 - 0.4)^{0} = 0.064 = 6.4\%$$

Ejemplo 2: Un lote de 1000 pernos contiene 15% de defectuosos. Cuál es la probabilidad de que al extraer 3 pernos al azar del lote se obtenga 1 defectuoso?

Objetivo del Ejemplo: Mostrar el cumplimiento de las condiciones de un experimento binomial y la aplicación de la distribución binomial.

Antes de abordar la solución del problema, observe que el espacio muestral de todas las muestras diferentes de 3 pernos (defectuoso = d y bueno = b) es: S = {(bbb), (dbb), (bdb), (bbd), (bdd), (dbd), (ddb), (ddd)}.

La probabilidad de ocurrencia de cada uno de los elementos del espacio muestral considerando las condiciones de la definición (P(d) = 0.15 y P(b) = 0.85: como invariables y que el resultado de defectuoso o bueno de cualquier perno es independiente del resultado de cualquier otro), se muestra en la siguiente tabla:

Bbb	dbb	bdb	Bbd	bdd	dbd	ddb	ddd
$.85^3$	$.15x.85^2$	$.15x.85^2$	$.15x.85^2$	$.85x.15^2$	$.85x.15^2$	$.85x.15^2$	$.15^3$

Tabla 4.1. - Probabilidades de ocurrencia de los elementos del espacio muestral del ejemplo2

Solución:

Dado que la variable aleatoria de interés es el número de defectuosos obtenido en la muestra y no el orden en que "salen" , se tiene:

Nº Defectuosos	Elementos	Probabilidad	Valor
0	(bbb)	0.85^3	0.6141
1	(dbb),(bdb),(bbd)	$3x0.85^2 x0.15$	0.3251
2	(ddb),(bdd),(dbd)	$3x0.15^2 x0.85$	0.0574
3	(ddd)	0.15^3	0.0034

Tabla 4.2. - Distribución de probabilidad de la variable aleatoria del ejemplo 2

Retomando el problema la solución pedida es 32.51%

Para la aplicación de la formula binomial se tiene:

- El número de repeticiones es 3, es decir n = 3.
- En cada una de las repeticiones el resultado es defectuoso o no defectuoso.
- La probabilidad de que el perno extraído sea defectuoso en cada una de las 3 repeticiones puede considerarse constante, es decir, la probabilidad de que el perno sea defectuoso es del 15% en cualquiera de las tres extracciones.

Observe que esta suposición no se cumple de manera estricta en el problema, puesto que: P(primer perno sea defectuoso) = 150/1000 =0.15, mientras que P(2º defectuoso / 1º fue defectuoso) = 149/999 = 0.1492 y P(segundo defectuoso / primero fue bueno) = 150/999 = 0.1501.

Por lo tanto, la variable aleatoria X se distribuye binomialmente con:

$$\left.\begin{array}{l} n = 3 \\ p = 0.15 \\ k = 1 \end{array}\right\} \quad \text{Denotada como: } b(k; n, p) = b(1; 3, 0.15)$$

La probabilidad pedida es: $b(1; 3, 0.15) = {}_3C_1 \times (0.15)^1 \times (0.85)^2 = 32.51\%$

Como ilustración adicional las probabilidades de sacar 0, 1, 2, 3 pernos defectuosos y las probabilidades acumuladas se muestran en la tabla 4.3:

B(k; 3, 0.15)	${}_3C_k\ (0.15)^k\ (0.85)^{n-k}$	PROBABILIDAD	PROBABILIDAD ACUMULADA
B(0; 3, 0.15)	$({}_3C_0)\ (0.15)^0\ (0.85)^3$	61.41%	61.41%
B(1; 3, 0.15)	$({}_3C_1)\ (0.15)^1\ (0.85)^2$	32.51%	93.92%
B(2; 3, 0.15)	$({}_3C_2)\ (0.15)^2\ (0.85)^1$	5.74%	99.66%
B(3; 3, 0.15)	$({}_3C_3)\ (0.15)^3\ (0.85)^0$	0.34%	100.00%

Tabla 4.3. - Distribución de probabilidad de la variable aleatoria binomial

Observe que la columna probabilidad acumulada representa la probabilidad de obtener hasta k pernos defectuosos. Así, la probabilidad de obtener 0 o 1 defectuosos es del 93.92%.

Los valores de las probabilidades de variables aleatorias binomiales pueden hallarse directamente en tablas calculadas para diferentes valores de n, p y k.

MEDIA Y VARIANZA: La media μ y la varianza σ^2 de una variable aleatoria distribuida binomialmente, b(n, p), está dada por:

$$\mu = np \qquad \sigma^2 = npq$$

Ejemplo 3: Un lote de muchos pernos contiene una fracción defectuosa del 5%. Si se define la variable aleatoria X, como el número de pernos defectuosos obtenidos al extraer 25 de éstos del lote. Cuál es la media y la desviación estándar de X?

Objetivo del Ejemplo: Mostrar el calculo de la media y la varianza de una variable aleatoria distribuida binomialmente.

Solución:

$\mu = np = 25 \times 0.15 = 1.25$ $\sigma^2 = npq = 25 \times 0.05 \times 0.95 = 1.1875.$ $\sigma = 1.09$

2. DISTRIBUCIÓN HIPERGEOMETRICA

Múltiples problemas de control estadístico de calidad donde se pide la probabilidad de obtener k artículos defectuosos en una muestra de tamaño n extraída de un lote finito con una fracción defectuosa p, se asumían como experimentos binomiales pese a que la condición de que la probabilidad de defectuoso no permanecía constante en las sucesivas repeticiones del experimento (extracciones).

Sin embargo esta aproximación no es válida en muchos experimentos donde el tamaño del lote es pequeño y la relación entre los tamaños de la muestra y del lote es mayor al 10%. En estos casos se requiere el uso de la distribución hipergeométrica.

Considere un lote de N productos de los cuales k son defectuosos. Se extraen al azar n artículos del lote y se define la variable aleatoria X = x como el número de defectuosos obtenidos en la muestra. Bajo estas condiciones X se define como una variable aleatoria hipergeométrica y denotada como: h(x; N, n, k).

Observe que el número de maneras de obtener x defectuosos de los k artículos defectuosos del lote esta dado por: ${}_{K}C_{x}$. Y por cada una de estas maneras se pueden obtener ${}_{N-K}C_{n-x}$, artículos buenos, por lo tanto, el número total de maneras es: $({}_{K}C_{x})\ ({}_{N-K}C_{n-x})$ y representa el número de casos favorables. Ahora el número total de maneras como pueden seleccionarse muestras de tamaño n del lote de N esta dado por: ${}_{N}C_{n}$.

Por lo tanto, la probabilidad de obtener x defectuosos en la muestra de n artículos extraída del lote de tamaño N, de los cuales k son defectuosos y N – k son buenos, está dada, por:

$$h(x; N, n, k) = \frac{({}_{K}C_{x})\ ({}_{N-K}C_{n-x})}{{}_{N}C_{n}}$$

Media y varianza: La media y la varianza de la distribución hipergeométrica Estan dadas, por:

$$\mu = \frac{nk}{N} \quad y \quad \sigma^2 = \frac{(N-n)}{(N-1)}\ \mu\ \frac{(N-k)}{N}$$

Ejemplo 4: Se tiene un lote de 1000 pernos de los cuales 150 son defectuosos y 850 son buenos. Cuál es la probabilidad de obtener 1 defectuoso en una muestra de 4 pernos extraída al azar del lote?

Objetivo del ejemplo: Mostrar la aplicación de la distribución hipergeométrica y su diferencia con la distribución binomial.

Solución: Para este caso se tiene:

$$N = 1000 \qquad n = 4 \qquad k = 150 \qquad x = 1$$

$$h(1; 1000, 4, 150) = \frac{({}_{150}C_{1})\ ({}_{850}C_{3})}{{}_{1000}C_{4}} = 0.3694$$

A manera de comparación si se considera la variable aleatoria X distribuida binomialmente la probabilidad pedida es:

$$b(1; 4, 0.15) = 0.3685$$

Observe que la diferencia de ambos resultados no es muy significativa.

Ejemplo 5: En una caja hay 10 bolas de las cuales 4 son rojas y 6 son azules, si se extraen 3 bolas al azar, cuál es la probabilidad de obtener 1 bola roja? Cuál es el valor esperado del número de bolas rojas?

Objetivo del ejemplo: Mostrar la aplicación de la distribución hipergeométrica y su diferencia con la distribución binomial.

Solución: En este caso se tiene:

$$h(1;\,10,\,3,\,4) = \frac{({}_4C_1)\,({}_6C_2)}{{}_{10}C_3} = 0.5$$

El resultado considerando la variable aleatoria distribuida binomialmente es:

$$b(1;\,3,\,0.4) = 0.432$$

Observe que en este caso la diferencia de los resultados es significativa y por lo tanto no es aceptable asumir el experimento como del tipo binomial (La relación del tamaño de la muestra y el tamaño del lote es 3/10 = 30%)

El valor esperado de la variable aleatoria esta dado por:

$$E(X) = \sum_{i=0}^{3} x_i\, p(x_i)$$

Las probabilidades para X = 0, 1, 2 y 3 son:

$$\begin{aligned} P(X = 0) &= h(0;\,10,\,3,\,4) = 0.17 \\ P(X = 1) &= h(1;\,10,\,3,\,4) = 0.50 \\ P(X = 2) &= h(2;\,10,\,3,\,4) = 0.30 \\ P(X = 3) &= h(3;\,10,\,3,\,4) = 0.03 \end{aligned}$$

Por lo tanto, E(X) = 0x0.17 + 1x0.50 + 2x0.30 + 3x0.03 = 1.2; observe que este valor coincide con el valor dado para la media de la distribución hipergeométrica $\mu = (3x4) / 10 = 1.2$.

3. DISTRIBUCIÓN DE POISSON

Sea la variable aleatoria X que puede tomar los valores k = 0, 1, 2, ..., n, y la probabilidad de que X sea igual a k, está dada por:

$$P(X = k) = \frac{e^{-\alpha}\alpha^{k}}{k!} \quad \text{para } k = 1, 2, \ldots\ldots, n$$

Se define como una variable aleatoria distribuida Poisson con media α.

La validez de la definición anterior depende del cumplimiento de las siguientes condiciones:

- El número promedio de ocurrencias del suceso en un intervalo de tiempo o "región" es conocido y se denota como α.
- La probabilidad de ocurrencia del suceso es proporcional al tamaño del intervalo o región y no depende del número de ocurrencias por fuera del intervalo o región.
- La probabilidad de que el suceso ocurra más de una vez en un intervalo o región muy pequeño es despreciable.

Como para cualquier distribución de probabilidad debe cumplirse:

$$\sum_{K=0}^{\infty} P(X=k) = \sum_{K=0}^{\infty} e^{-\alpha}\alpha^{k} / k! = 1$$

Ejemplo 6: Un equipo de fútbol hace en promedio 2.2 goles por partido. ¿Cuál es la probabilidad de que en el próximo partido convierta 4 goles?

Objetivo del ejemplo: Mostrar la aplicación de la distribución Poisson.

Solución:

$$\alpha = 2.2. \quad P(X = 4) = e^{-2.2}\, 2.2^{4} / 4! = 10.82\%$$

La siguiente tabla muestra las probabilidades de que el equipo haga 0, 1, 2,.... goles en el próximo partido, así como las probabilidades acumuladas.

Nº de goles	Probabilidad	Probabilidad acumulada
0	11.08%	11.08%
1	24.38%	35.46%
2	26.81%	62.27%
3	19.66%	81.93%
4	10.82%	92.75%
5	4.75%	97.50%
6	1.74%	99.24%
7	0.55%	99.79%
8	0.15%	99.94%

Tabla 4.4. - Distribución de probabilidad de la variable aleatoria Poisson del ejemplo 6

Ejemplo 7: Registros históricos muestran que una central eléctrica sale de servicio 7 veces por año. Cuál es la probabilidad de que en un año cualquiera haya menos de 3 cortes del servicio?

Objetivo del ejemplo: Mostrar la aplicación de la distribución Poisson y el calculo de probabilidades acumuladas.

Solución:

La probabilidad pedida es: $\sum_{K=0}^{2} P(X = k) = \sum_{K=0}^{2} e^{-7} 7^{k} / k! = 2.97\%$

Los valores de las probabilidades acumuladas de la distribución Poisson pueden obtenerse de tablas calculadas para diferentes valores de α y k.

4. APROXIMACIONES ENTRE LAS DISTRIBUCIONES

En general, los “experimentos” usados por el control estadístico de calidad, consisten en examinar una característica de calidad a muestras provenientes de lotes de productos sin sustitución.

Esto nos lleva a considerar el uso de la distribución hipergeométrica para determinar las probabilidades de hallar en la muestra un número dado de artículos defectuosos si el lote de la cual proviene tiene una fracción defectuosa determinada.

Ejemplo 8: Como ilustración considere un lote de 50 ejes que pueden clasificarse como defectuosos o como buenos. Suponga que en el lote hay 10 ejes defectuosos y 40 buenos. Si se toma de este lote una muestra de 5 ejes, cuál es la probabilidad de que haya un defectuoso?

Objetivo del ejemplo: Mostrar como las condiciones del experimento determinan el uso de las distribuciones de probabilidad.

Solución:

Observe que en este experimento la extracción de los 5 ejes se realiza sin reposición y por lo tanto las probabilidades de que sucesivos ejes sean defectuosos depende de los resultados obtenidos en las extracciones previas y por lo tanto el experimento se ajusta estrictamente a la distribución hipergeometrica, donde: $N = 50$; $n = 5$; $x = 1$; $k = 10$

Por lo tanto, la probabilidad pedida es: $P(x = 1) = h(1; 50; 5; 10) = 0.431$

El uso de la distribución binomial para este caso implica comprobar si el experimento cumple las condiciones establecidas para esta distribución, es decir que la probabilidad de hallar un eje defectuoso en las sucesivas extracciones permanezca constante. Evidentemente esto no se cumple, puesto que estas probabilidades cambian dependiendo de los resultados obtenidos en las extracciones previas.

Observe, por ejemplo, que para la primera extracción la P(1º defectuoso) = 10/50, mientras que para la segunda extracción esta probabilidad depende de si en la primera extracción el eje fue defectuoso o bueno, así: P(2º defectuoso / 1º fue defectuoso) = 9/49, mientras que la P(2º defectuoso / 1º fue bueno) = 10/49.

Pese a esta circunstancia en muchas ocasiones se puede considerar que las probabilidades son aproximadamente iguales para las sucesivas extracciones y en consecuencia puede usarse la distribución binomial.

Para el caso se tiene: $P(x = 1) \approx b(1; 5; 0.20) = 0.4096$, donde, $p = N / k = 10/50 = 0.20$.

Como puede verse la diferencia del cálculo mediante ambas distribuciones es significativa (43.13% y 40.96%)

- **LA DISTRUCIÓN BINOMIAL COMO APROXIMACIÓN DE LA HIPERGEOMÉTRICA**

Dependiendo del grado de precisión requerido puede usarse la distribución binomial como una aproximación de la distribución hipergeométrica, cuando la relación entre el tamaño de la muestra y el tamaño del lote es menor del 5%, obteniéndose resultados “razonablemente” similares.

Ejemplo 9: Un lote de 5000 rodamientos contiene 250 defectuosos. Cuál es la probabilidad de no obtener defectuosos al tomar una muestra de 10 de estos rodamientos?

Objetivo del ejemplo: Mostrar las condiciones requeridas para usar la distribución binomial como aproximación de la distribución hipergeométrica.

Solución:

Usando la distribución hipergeometrica se tiene:

$$P(x = 0) = h\ (0;\ 5000;\ 10;\ 250) = 59.85\%$$

Usando la distribución binomial con $p = 250/5000 = 0.05$

$$P(x = 0) = b(0;\ 10;\ 0.05) = 59.87\%$$

Observe que ambos resultados son muy similares, debido a que la relación entre el tamaño de la muestra y el tamaño del lote es del 0.2%.

- **LA DISTRIBUCIÓN POISSON COMO APROXIMACIÓN DE LA DISTRIBUCIÓN BINOMIAL**

El uso de la distribución Poisson como aproximación de la distribución binomial puede hacerse cuando el tamaño de la muestra es grande y p es cercano a cero. Los resultados son aproximadamente iguales para valores de $n \geq 50$ y $p \leq 5\%$. Donde el parámetro $\alpha = np$

Ejemplo 10: Un lote de remaches tiene una fracción defectuosa del 1%, Cuál es la probabilidad de no obtener defectuosos en una muestra de 100?

Objetivo del ejemplo: Mostrar las condiciones requeridas para usar la distribución Poisson como aproximación de la distribución binomial.

Solución:

Usando la distribución binomial: $P(x = 0) = b(0; 100; 0.01) = 36.61\%$

Usando Poisson con $\alpha = np = 100 \times 0.01 = 1$: $P(x = 0) = 36.78\%$

Observe que ambos resultados son muy similares, debido a que el tamaño de la muestra es grande y p cercano a cero.

5. TALLER

1. - Se tienen las etiquetas de 2 bolsas de maíz pira (para crispetas):

MARCA A	MARCA B
Contenido: 100 granos GARANTIA: Cada grano de esta bolsa será una crispeta 99 de cada 100 veces	Contenido: 100 granos GARANTIA: 99 de los 100 granos de esta bolsa serán crispetas.

- Interprete las garantías de ambas marcas.
- Determine para las bolsas de cada marca las probabilidades de que no “estallen” 0, 1, ... , k granos.
- Calcule e interprete el valor esperado (media) de ambas marcas

2. En relación con el problema 1.

- Calcule e interprete las probabilidades de que no estallen 0, 1, 2 y 3 granos de una bolsa de la marca A.
- Calcule e interprete las probabilidades de que no estallen 0, 1, 2 y 3 granos preparando media bolsa de la marca B.

3. En un cajón hay 10 bolas de las cuales 3 son rojas y el resto blancas. Si se sacan tres al azar cual es la probabilidad de que todas sean rojas?

4. En las siguientes situaciones calcule la probabilidad de encontrar hasta 1 defectuoso y justifique la distribución que usó:

- Lote = 5; con 3 defectuosos en el lote; Muestra = 2
- Lote = 100; con 3 defectuosos en el lote; Muestra = 10
- Lote = 100; con 30 defectuosos en el lote; Muestra = 10

5. En la Feria del Calzado hay un contenedor con muchos zapatos del mismo estilo, color y talla, de los cuales el 60% son derechos y el 40% izquierdos. Un concursante elige al azar 10 zapatos del contenedor ganando todos los pares (derecho – izquierdo) que saque. Calcule las probabilidades de:

- De ganarse 0 pares
- De ganarse 1 par
- De ganarse 2 pares
- De ganarse 3 pares
- De ganarse 4 pares
- De ganarse 5 pares

6. Un estudiante se sometió a un test de 10 preguntas de escogencia múltiple con 4 alternativas. Pese a haber estudiado "mucho se le fue la paloma" y contesto las preguntas al azar. Del resultado de la prueba depende la permanencia del estudiante en la Universidad, quedando por fuera si saca malas más de la mitad las preguntas.

- Analice las condiciones de este "experimento"
- Que distribución de probabilidad se adapta a estas condiciones?
- Cuál es la probabilidad de que el estudiante no salga de la U?

7. Un telar produce un tejido con un promedio de dos defectos por metro cuadrado. Cuál es la probabilidad de que haya 4 defectos en un metro ^{2}.

8. Considere el experimento de lanzar tres veces un dado.

- Cuál es la probabilidad de obtener tres "unos". (Justifique la distribución de probabilidad usada para calcular esta probabilidad)
- Calcule esta probabilidad usando la distribución Poisson.
- Discuta la validez del uso de la distribución Poisson para la solución del problema

9. La producción de una máquina es 20% defectuosa. Se seleccionan al azar 10 artículos producidos por esta máquina.

- Cuál es la probabilidad de que 9 o más artículos estén buenos. Use las distribuciones Binomial y Poisson y compare las respuestas.
- Si la fracción defectuosa es del 30% ¿cómo cambia la solución?
- Si la fracción defectuosa es del 10% ¿cómo cambia la solución?
- Qué puede concluirse de las soluciones de los tres puntos anteriores?

10. La media y la desviación estándar de la variable aleatoria X, distribuida binomialmente, son: 18.75 y 3.75, respectivamente. Calcule los parámetros n y p de la distribución e interprete los resultados.

11. Un cazador de conejos acierta el 82% de sus disparos. Cada mañana sale con cinco flechas a buscar el sustento de su familia. ¿Cuántos conejos se espera que cace en los siguientes 30 días?

6. - BIBLIOGRAFÍA

- Meyer, P. *Probabilidad y Aplicaciones Estadísticas.* Fondo Educativo Interamericano. México 1981.

- Ash, R. *Basic Probability Theory.* John Wiley & Sons. New York. 1976.

- Walpole, R. *Introduction to Statistics.* Macmillan Company. New York 1969.

- Mason, R. y D. Lind. *Estadística para Administración y Economía* Alfaomega. México 1992

MODULO 5 LA DISTRIBUCIÓN NORMAL

OBJETIVOS DEL MODULO:

- Comprender la naturaleza de las distribuciones continuas de probabilidad
- Comprender las características y usos de la Distribución Normal

CONTENIDO:

PRERREQUISITOS:

- Módulos 2, 3 y 4

1. DISTRIBUCIONES CONTINUAS DE PROBABILIDAD

Las distribuciones binomial, hipergeométrica y poisson son discretas, es decir, las variables aleatorias sólo pueden tomar valores discretos (tal como número de unidades defectuosas).

Las variables aleatorias continuas, en cambio, pueden tomar todos los valores posibles en un intervalo dado y por lo tanto no es posible su "enumeración" debido a que su número es infinito. A manera de ilustración considere la variable aleatoria X como el peso de los alfileres producidos por una máquina. Asuma que el peso de los alfileres está en el rango (25 – 35) miligramos. Evidentemente no sería posible identificar el número de pesos diferentes en un lote de estos alfileres, por ejemplo si se dice que hay alfileres de 25, 25.1, 25.2, 25.3, ..., 34.8, 34.9 y 35 miligramos, donde se clasificaría un alfiler que pese 25.11 miligramos?

Lo anterior implica que no es posible asignar la probabilidad de ocurrencia para "un valor dado" de la variable aleatoria y en consecuencia la probabilidad de este evento es cero.
Observe que si en la variable aleatoria considerada la probabilidad de que los alfileres estén entre 25 y 35 miligramos es uno, $P(25 \leq X \leq 35) = 1$.

Además, $P(25 \leq X \leq 35) = P(X = 25) + P(25 < X < 35) + P(X = 35) = P(25 < X < 35)$, dado que $P(X = 25)$ y $P(X = 35)$ valen cero.

Por lo tanto, las probabilidades asociadas a la variable aleatoria deberán representarse mediante una función continua entre los límites del rango o intervalo de ésta.

2. LA DISTRIBUCIÓN NORMAL

La distribución normal es la distribución de probabilidad continua más importante en el campo de la estadística. La función de probabilidad de la distribución normal es de forma acampanada y se conoce como curva normal.

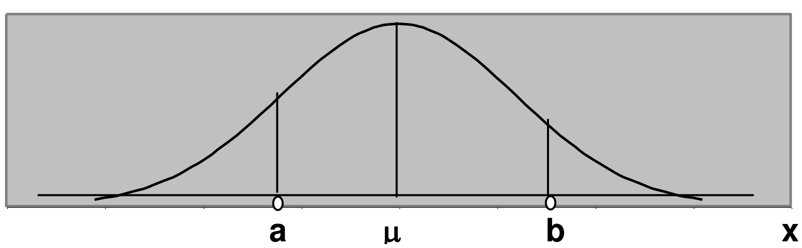


Figura 5.1. - Distribución de probabilidad Normal.

Si la función de probabilidad de la variable aleatoria X, **f**(X), tiene la forma acampanada como la mostrada en la figura, se dice que X se distribuye normalmente. La distribución normal tiene los parámetros μ y σ, los cuales representan la media y la desviación estándar.

La probabilidad $P(a < X < b)$ = Area determinada entre los valores a y b del eje x y la función de probabilidad, Por lo tanto, $P(a < X < b) = \int_a^b f(X)dx$

La curva normal tiene las siguientes características:

- Es simétrica respecto a la ordenada que pasa por la media μ..
- Es asintótica a los extremos del eje x.
- El área total bajo la curva es 1.
- El valor σ determina el grado de "apuntamiento" o "achatamiento" de la campana.

3. **LA CURVA NORMAL ESTANDAR:** Es posible transformar cualquier curva normal a una curva normal equivalente con parámetros $\mu = 0$ y $\sigma = 1$, mediante la definición de la variable aleatoria $Z = (X - \mu) / \sigma$, y llamada normal estándar o normal (0, 1).

Ejemplo 1: Retomando el caso de los alfileres y asumiendo que el peso se distribuye normal con $\mu = 30$ miligramos y $\sigma^2 = 4$. Cuál es la probabilidad de que un alfiler pese 29 o menos miligramos?

Objetivo del ejemplo: Mostrar la aplicación de la distribución normal, su estandarización y el uso de la tabla normal (0 – 1)

Solución:

$P(X \leq 29)$, es equivalente a $P\big((Z \leq (29 - 30) / 2\big) = P(Z \leq -0.5)$.

Los valores bajo la curva normal (0, 1) pueden hallarse en una tabla de áreas normal (0, 1), ubicando el valor correspondiente de z. (Ver tabla)

Para este caso, con z = -0.5, puede leerse en la tabla de áreas normal (0, 1) el valor 0.3085, es decir la probabilidad de que un alfiler pese menos de 29 miligramos es del 30.85%.

Ejemplo 2: El ingreso promedio de los habitantes de un país se distribuye normal con una media U$ 400 y desviación estándar de 60. Cuál es la probabilidad de un ciudadano tenga un ingreso de más de U$ 510?

Objetivo del ejemplo: Mostrar la aplicación de la distribución normal, su estandarización y el uso de la tabla normal (0 – 1) o estándar.

Solución:

$P(X > 510) = 1 - P(X \leq 510) = 1 - P(Z \leq (510 - 400) / 60) = 1 - P(Z \leq 1.833)$

$P(X > 510) = 1 - 0.9664 = 0.0336 = 3.36\%$

Ejemplo 3: Basado en los datos del problema anterior. Se quiere saber cuál es el valor tope del ingreso que determina al 15% de los habitantes de más bajos ingresos?

Objetivo del ejemplo: Mostrar la aplicación de la distribución normal estándar y el uso de la tabla normal (0 – 1), para hallar el valor de la variable aleatoria correspondiente a una probabilidad dada.

Solución:

En este caso, se pide hallar el valor de x, para el cual, $P(X \leq x) = 0.15$, o de manera equivalente $P(Z \leq (x - 400)/60) = 0.15$. En la tabla puede leerse que para un área de 0.1492 el valor correspondiente de $z = -1.04$, por lo tanto, $x = (-1.04 \times 60) + 400 = 337.6$ dólares.

4. LA DISTRIBUCIÓN NORMAL COMO APROXIMACIÓN DE LA BINOMIAL

Para valores de n grandes y p y q no cercanos a cero puede usarse la distribución normal como aproximación de la distribución binomial, con media, $\mu = np$ y $\sigma = \sqrt{npq}$. En términos generales la aproximación es buena cuando simultáneamente se cumplen que np y nq sean mayores que 5.

Dado que la distribución binomial es discreta y la distribución normal es continua, es necesario "ajustar" el valor de la variable aleatoria binomial, así:
La probabilidad de que la variable aleatoria binomial, X, esté entre x_k y x_j (ambos valores enteros), en n repeticiones, está dada por: $\sum_{i=k}^{i=j} b(x_i; n; p)$

Para el cálculo de esta misma probabilidad usando la distribución normal, se tiene: $P(x_k - 0.5 \leq X \leq x_J + 0.5)$, los valores de ± 0.5, tienen como efecto el "darle" continuidad a la variable aleatoria original que es discreta.

Ejemplo 4: Calcular la probabilidad de obtener 3 o 4 caras al lanzar una moneda 10 veces, usando la distribución binomial y la distribución normal.

Objetivo del ejemplo: Mostrar la aplicación de la distribución normal, como aproximación de la distribución binomial.

Solución:

Usando la distribución binomial, se tiene: b(3; 10; 0.5) + b(4; 10, 0.5) = 32.23%

Usando la distribución normal, con μ = n x p = 5 y σ = $(nxpxq)^{0.5}$ = 1.58, se tiene: P(3 – 0.5 $\leq$ X $\leq$ 4 + 0.5) = P(- 2.5 / 1.58 $\leq$ Z $\leq$ - 0.5 / 1.58) = 0.3745 – 0.0571 = 31.74%

Observe que ambos resultados están cercanos

Ejemplo 5: Calcular la probabilidad de obtener 20 o menos caras al lanzar una moneda 50 veces, usando la distribución binomial y la distribución normal.

Objetivo del ejemplo: Mostrar como mejora la aproximación entre las distribuciones al incrementar n.

Solución:

Usando la distribución binomial: $\sum_{K=0}^{K=20} b(k;50;0.5)$ = 10.13%

Usando la distribución normal, con μ = n x p = 25 y σ = $(n \times p \times q)^{0.5}$ = 3.535, se tiene: P(0 - 0.5 $\leq$ X $\leq$ 20 + 0.5) = P(- 25.5 / 3.535 $\leq$ Z $\leq$ - 4.5 / 3.535) = 0.1020 – 0 = 10.2%

Ambos resultados muy cercanos.

5. TALLER

1. El nivel de glucosa en la sangre es una variable aleatoria distribuida normalmente con μ = 105 mg. / cc. y σ = 15 mg. / cc. Si la Organización Mundial de la Salud estableció el rango { 80 – 120 } como valores de referencia indicadores de buena salud. ¿Qué porcentaje de la población estará por fuera de este rango?

2. La nota promedio de los 3000 estudiantes de la Facultad se distribuye normal. Si el pasado semestre la media fue de 3.2 y la desviación estándar 0.4.

- Cuántos estudiantes obtuvieron promedio de 4.0 o más?
- Cuántos estudiantes obtuvieron promedio de 2.0 o menos?
- Cuántos estudiantes obtuvieron promedio entre 2.5 y 3.5?
- Si la universidad concede a los 150 mejores estudiantes exención de matricula que promedio como mínimo debió obtenerse para merecer este reconocimiento?
- Si el semestre antepasado la media fue igual y la cuarta parte de los estudiantes sacaron 3.5 o más, cuál fue el valor de la desviación estándar durante este semestre?

3. Una máquina produce dados de acero cuyo peso se distribuye normalmente con media igual a 170 gramos y desviación estándar 10 gramos. Los dados se consideran buenos si pesan más de 169 y menos de 173 gramos.

- Cuál es la probabilidad de que un dado sea defectuoso?
- Si se producen dos de estos dados ¿Cuál es la probabilidad de que uno sea bueno y el otro defectuoso?

4. Una empresa renta baterías recargables por día (24 horas). La duración de éstas en horas se distribuye N(20; 2). La tarifa del alquiler por hora depende del tiempo efectivo de funcionamiento, así: Si la batería dura menos de 18 horas no se cobra la tarifa al cliente. Si la batería opera más de 18 pero menos de 20 horas la tarifa es de \$ 1500 / día. Si la batería opera más de 20 horas la tarifa es de \$ 2000 / día.

- Cuál es el ingreso promedio de la empresa por batería – día rentada?

5. La duración en horas de un tipo de bombillo se distribuye normal con media de 1000 horas. Si la probabilidad de éstos duren más de 942 horas es del 98%, cuál es la desviación estándar de la duración de estos bombillos?

6. Una máquina dispensadora de café, sirve raciones distribuidas normalmente, con μ = 35cc y σ = 4 cc. Si las tasas usadas tienen una capacidad de 40 cc, cuál es la probabilidad de que haya un derrame de la bebida?

7. Se tienen dos marcas de baterías cuya duración se distribuye normal. Si la batería A dura en promedio 750 horas y la desviación estándar es de 45 horas y la batería B dura en promedio 780 horas y la desviación estándar es de 18 horas. Cuál de las baterías compraría, si se desea que la batería dure más de 800 horas?

8. La longitud de los pasadores producidos por las máquinas A y B, se distribuye normal. Ambas máquinas producen con μ = 18 cm y se sabe que la desviación estándar de la producción de la máquina A es σ = 1.5 cm.

- Cuál es la probabilidad de que un pasador producido por la máquina A tenga una longitud menor de 17 cm?
- Cuál es la desviación estándar de la longitud de los pasadores producidos por la máquina B, si la probabilidad de que un pasador sea menor de 17 cm es la mitad que la de los producidos por A?

9. La estatura en metros de los habitantes de una comunidad de 25000 personas se distribuye N(1.74, 0.12).

- Cuál es la probabilidad de un habitante elegido al azar tenga una estatura comprendida en el rango de la media $\pm$ 1 desviación estándar?
- Cuántos habitantes de la comunidad tendrán una estatura comprendida en el rango de la media $\pm$ 2 desviaciones estándar?
- Cuántos habitantes tendrán una estatura por fuera del rango de la media $\pm$ 3 desviaciones estándar?

10. Un contratista sabe que el tiempo de construcción de edificios se distribuye normal. Si acaba de aceptar un contrato de construcción, para la cual estimó un tiempo de construcción promedio de 42 meses y desviación estándar de 2.5 meses, y el contrato establece que la obra debe entregarse como máximo en 38 meses. Cuál es la probabilidad de que pueda entregar la obra en el tiempo estipulado?

6. TABLA NORMAL ESTANDAR

	0,00	**0,01**	**0,02**	**0,03**	**0,04**	**0,05**	**0,06**	**0,07**	**0,08**	**0,09**
-3,1	0,0010	0,0009	0,0009	0,0009	0,0008	0,0008	0,0008	0,0008	0,0007	0,0007
-3,0	0,0013	0,0013	0,0013	0,0012	0,0012	0,0011	0,0011	0,0011	0,0010	0,0010
-2,9	0,0019	0,0018	0,0018	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014
-2,8	0,0026	0,0025	0,0024	0,0023	0,0023	0,0022	0,0021	0,0021	0,0020	0,0019
-2,7	0,0035	0,0034	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026
-2,6	0,0047	0,0045	0,0044	0,0043	0,0041	0,0040	0,0039	0,0038	0,0037	0,0036
-2,5	0,0062	0,0060	0,0059	0,0057	0,0055	0,0054	0,0052	0,0051	0,0049	0,0048
-2,4	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064
-2,3	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084
-2,2	0,0139	0,0136	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0110
-2,1	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143
-2,0	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183
-1,9	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233
-1,8	0,0359	0,0351	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0301	0,0294
-1,7	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367
-1,6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455
-1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559
-1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0721	0,0708	0,0694	0,0681
-1,3	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823
-1,2	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985
-1,1	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170
-1,0	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379
-0,9	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
-0,8	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867
-0,7	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2177	0,2148
-0,6	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451
-0,5	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2843	0,2810	0,2776
-0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
-0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483
-0,2	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
-0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
-0,0	0,5000	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993

7. BIBLIOGRAFÍA

- Meyer, P. *Probabilidad y Aplicaciones Estadísticas.* Fondo Educativo Interamericano. México 1981.

- Ash, R. *Basic Probability Theory.* John Wiley & Sons. New York. 1976.

- Walpole, R. *Introduction to Statistics.* Macmillan Company.New York 1969.

- Mason, R. y D. Lind. *Estadística para Administración y Economía.* Alfaomega. México 1992

- Mood, A. y F. Graybill. *Introducción a la Teoría de la Estadística.* Aguilar. Madrid. 1972

- Shao Stephen. *Estadística para Economistas y Administradores de Empresas.* Herrero Hermanos. México. 1979.

MODULO 6 TEORIA DEL MUESTREO

OBJETIVOS DEL MODULO:

- Comprender la Teoría del Muestreo
- Comprender la relación entre poblaciones y muestras aleatorias
- Conocer los principales estadísticos de muestras
- Comprender los conceptos y usos de las distribuciones de muestreo

CONTENIDO:

PRERREQUISITOS:

- Módulos 4 y 5

1. TEORIA DEL MUESTREO

La teoría del muestreo trata sobre las relaciones existentes entre los valores que miden algunas características de una población, (llamados parámetros poblacionales), y los valores obtenidos de estas características a partir de observaciones obtenidas en muestras extraídas de la población.

- **POBLACIÓN:** La población puede definirse como la totalidad de las observaciones del fenómeno de interés. El número de observaciones de la población está definido por el tamaño de la población, generalmente denotado como N, éste puede ser finito o infinito.

  Por ejemplo, el número de estudiantes matriculados en un semestre dado en la Sede genera N observaciones, caso en el cuál la población es finita. Por el contrario, si el fenómeno de interés es el número de puntos obtenidos al lanzar un dado indefinidamente, la totalidad de observaciones posibles constituye una población infinita.

- **PARÁMETRO:** Un parámetro poblacional se define como el valor de una característica dada de la población. Como ejemplos pueden citarse: la estatura promedio de los estudiantes matriculados en un semestre dado en la sede o la proporción de electores a favor de un candidato de la totalidad de electores de una comunidad.

- **MUESTRA:** Una muestra es un subconjunto de la población.

- **MUESTRA ALEATORIA:** Una muestra de tamaño n extraída de una población de tamaño N, es aleatoria si todos los subconjuntos de tamaño n tienen la misma probabilidad de ser extraídos de la población.

- **ESTADÍSTICO:** Es el valor de una característica, calculado a partir de los elementos de una muestra. Observe que este valor puede variar para cada una de las posibles muestras extraídas de la población, y, por lo tanto, este valor se constituye en una variable aleatoria.

2. ESTADÍSTICOS IMPORTANTES

Se definen a continuación los estadísticos más usados en el Control de Calidad, así:

- **MEDIA DE MUESTREO:** Si $x_1, x_2, x_3, ..., x_n$ representan valores de una característica observados en una muestra aleatoria de tamaño n, la media de muestreo, está dada por:

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n}$$

- **RANGO DE MUESTREO:** El rango R de los valores observados en una muestra aleatoria de tamaño n, $x_1, x_2, x_3, ..., x_n$, es la diferencia entre el mayor y el menor valor del conjunto.

$$R = \text{Max}\,[x_i] - \text{Min}\,[x_i] \quad \text{para } i = 1, 2, ..., n$$

- **VARIANZA DE MUESTREO:** Si $x_1, x_2, x_3, ..., x_n$ representan los valores de una característica observados en una muestra aleatoria de tamaño n, la varianza de muestreo está dada por:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

Una formula alternativa de la varianza de muestreo, que ofrece ventajas de cálculo, es:

$$s^2 = \frac{n\sum_{i=1}^{n} x_i^{\,2} - \left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)}$$

La desviación estándar de muestreo denotada como, s, es la raíz cuadrada de la varianza.

3. DISTRIBUCIONES DE MUESTREO

La aplicación de la estadística inductiva al campo del control estadístico de calidad permite inferir características de un lote (población) a partir de la observación de estas características en una muestra.

Dado que el valor de un estadístico de muestreo depende exclusivamente de la muestra, este se constituye en una variable aleatoria. Así, puede asignarse la probabilidad de ocurrencia de los valores del estadístico asociados a cada una de las posibles muestras que pueden extraerse de la población. Y esto se define como la distribución del muestreo de probabilidad del estadístico.

4. **DISTRIBUCIÓN MUESTRAL DE LA MEDIA:** Para el caso de la Media, se trata de establecer la distribución de probabilidad de la media obtenida de todas las posibles muestras aleatorias de tamaño n extraídas de la población.

Suponga que en una caja hay cuatro balotas, las cuales están marcadas con los dígitos 1, 2, 3 y 4. Si se define la variable aleatoria, X, como el valor de la marca de las balotas, la población (N = 4) tiene la siguiente distribución de probabilidad:

Valor de X	Probabilidad
1	0.25
2	0.25
3	0.25
4	0.25

Tabla 6.1. Distribución de probabilidad de la población

La media y la varianza de la población, son:

$$\mu = E(X) = 1x\ 0.25 + 2x0.25 + 3x0.25 + 4x0.25 = 2.5$$

$$\sigma^2 = \{(1-2.5)^2 + (2-2.5)^2 + (3-2.5)^2 + (4-2.5)^2\} / 4 = 5 / 4 = 1.25$$

Suponga la extracción de muestras de dos balotas (n = 2), las 16 muestras posibles de tamaño 2 extraídas de la caja con sustitución y el cálculo de las medias de cada una de estas 16 muestras (medias de muestreo), se dan en la siguiente tabla:

Muestra	1,1	1,2	1,3	1,4	2,1	2,2	2,3	2,4	3,1	3,2	3,3	3,4	4,1	4,2	4,3	4,4
$\overline{X}$	1	1.5	2	2.5	1.5	2	2.5	3	2	2.5	3	3.5	2.5	3	3.5	4

Tabla 6.2. Medias de las muestras (n = 2)

Ahora, el estadístico $\overline{X}$**,** es una variable aleatoria que toma valores entre 1 y 4, la distribución de muestreo de $\overline{X}$**,** se puede ver en la siguiente tabla:

Valor de $\overline{X}$	Nº de veces	Probabilidad
1	1	1/16
1.5	2	2/16
2	3	3/16
2.5	4	4/16
3	3	3/16
3.5	2	2/16
4	1	1/16

Tabla 6.3. Distribución de probabilidad de la media de las 16 muestras

Con media y varianza dadas por:

$$\mu_X = \sum \overline{x} f(\overline{x}) = \mu = 2.5$$

$$\sigma_x^2 = \Sigma(\overline{X} - \mu_x)^2 f(\overline{X}) = 2.5 = \sigma^2 / 2 = 0.625$$

Generalizando el anterior ejemplo, puede definirse el siguiente teorema:

Para todas las posibles muestras de tamaño n extraídas con sustitución de una población de tamaño N con media μ y desviación estándar σ, la distribución de muestreo de la media, $\overline{X}$, se distribuye aproximadamente normal con media $\mu_x = \mu$, y desviación estándar, $\sigma_x = \sigma / \sqrt{n}$

La validez de este teorema esta condicionada a que $n \geq 30$, sin embargo para muestras menores de 30 el teorema es válido sólo si la población se distribuye normal o aproximadamente normal.

Ejemplo 1: Retomando el caso de las balotas, cuál es la probabilidad de que una muestra de 40 balotas extraídas con sustitución tenga un media menor de 1.9?

Objetivo del ejemplo: Mostrar la relación entre los parámetros media y desviación estándar con los correspondientes estadísticos del muestreo y su aplicación en el calculo de probabilidades.

Solución:

Se conoce que la media y la varianza poblacional, son:

$$\mu = 2.5 \text{ y } \sigma^2 = 1.25$$

Aplicando el teorema, la media de la muestra se distribuye aproximadamente normal con: $\mu_X = \mu = 2.5$ y $\sigma_X^2 = \sigma^2/40 = 1.25/40 = 0.03125$

Por lo tanto, $P(\overline{X} \leq 1.9) = P(Z \leq ((1.9 - 2.5) / 0.1768) = P(Z \leq -3.39)$, el área bajo curva normal (0, 1) da la probabilidad pedida y en consecuencia $P(\overline{X} \leq 1.9) = 0.03\%$.

Para el caso de muestras sin sustitución de tamaño mayor o igual a 30, el citado teorema se cumple sólo si el tamaño de la población, N, es grande o infinito.

Ejemplo 2: Se sabe que la media de las asignaturas registradas en un semestre dado por los estudiantes de la sede es de 3.5, con una desviación estándar de 2.5. Cuál es la probabilidad de que una muestra de 30 estudiantes tomada al azar tenga entre 4 y 5 asignaturas registradas?

Objetivo del ejemplo: Mostrar la relación entre los parámetros media y desviación estándar con los correspondientes estadísticos del muestreo y su aplicación en el calculo de probabilidades.

Solución:

Se tiene:

$$\mu = 3.5 \text{ y } \sigma = 2.5$$

La distribución del muestreo de la media para n = 30, es aproximadamente normal con:

$$\mu_X = \mu = 3.5 \text{ y } \sigma_X = \sigma / \sqrt{n} = 2.5 / \sqrt{30} = 0.4564$$

La probabilidad pedida es: $P(4 \leq \overline{X} \leq 5) = P(\overline{X} \leq 5) - P(\overline{X} \leq 4) = P(Z \leq (5 - 3.5) / 0.4564) - P(Z \leq (4 - 3.5) / 0.4564) = P(Z \leq 3.29) - P(Z \leq 1.09) = 0.9995 - 0.8621 = 13.74\%$

En el caso general para poblaciones finitas de tamaño N, con media μ y desviación estándar σ, la distribución del muestreo de la media $\overline{X}$**,** para muestras de tamaño n sin sustitución, se distribuye aproximadamente normal con media y desviación estándar dadas por:

$$\mu_\xi = \mu \quad \text{y} \quad \sigma_\xi = (\sigma / \sqrt{n}) \sqrt{\frac{N - n}{N - 1}}$$

Observe que para valores de N muy grandes en relación a n, la expresión $\sqrt{\frac{N-n}{N-1}}$ tiende a 1 y por consiguiente σ_x se aproxima a σ / $\sqrt{n}$

Tal es el caso del ejemplo 2, donde N = 6900 (número de estudiantes de la Sede) y la muestra es de 30 estudiantes, por lo tanto:

$$\sqrt{\frac{N-n}{N-1}} = \sqrt{\frac{6900-30}{6900-1}} = 0.998$$

y σ_x se aproxima a σ / $\sqrt{n}$ = 2.5 / $\sqrt{30}$ = 0.4564

Ejemplo 3: El ingreso promedio de un grupo de 15 familias es de 1.45 millones de pesos con una desviación estándar de 0.3 millones. Cuál es la probabilidad de que una muestra de 4 familias tomadas al azar sin sustitución tenga un ingreso promedio menor de 1.2 millones?

Objetivo del ejemplo: Mostrar el caso general de la relación entre los parámetros media y desviación estándar con los correspondientes estadísticos de muestras y su aplicación en el calculo de probabilidades.

Solución:

$$\mu_X = 1.45 \qquad y \qquad \sigma_x = (\,0.3\ /\ \sqrt{4}\,)\sqrt{\frac{15-4}{15-1}} = 0.133$$

Por lo tanto, $P(\overline{X} \leq 1) = P(Z \leq (1.2 - 1.45)/0.133) = P(Z \leq -1.88) = 3.01\%$

5. DISTRIBUCIÓN DEL MUESTREO DE PROPORCIONES

Recordando que la distribución binomial, b(n;p,q) tiene media, μ = np y varianza σ^2 = n x p x q. Las muestras aleatorias de tamaño n, extraídas de una población binomial, con media, μ = n x p y varianza σ^2 = n x p x q, constituyen la distribución del muestreo de las proporciones, P, la cual se distribuye aproximadamente normal con media μ_p = p y desviación estándar, $\sigma_p = \sqrt{\frac{pq}{n}}$

La condición de la distribución binomial relativa a que p permanezca constante, implica considerar un factor de corrección en el valor de la varianza de la distribución del muestreo de las proporciones, así:

Para poblaciones finitas $\sigma_p = \sqrt{\frac{pq}{n} x \frac{N-n}{N-1}}$

Observe que el factor de corrección $\frac{N-n}{N-1}$, tiende a 1 cuando N es muy grande o cuando la relación n / N es pequeña.

Ejemplo 4: Una máquina de producción continua de ejes produce 5% de defectuosos. Cual es la probabilidad de que una muestra de 50 ejes tenga entre 0% y 2% de defectuosos?

Objetivo del ejemplo: Mostrar la relación entre la proporción poblacional y la proporción de muestras y su aplicación en el calculo de probabilidades.

Solución:

$\mu_p = p = 0.05$ y $\sigma_p = \sqrt{\frac{pq}{n}} = 0.031$

La probabilidad de que P esté entre 0% y 2%, $P(0.00 - 0.5/50 \leq P \leq 0.02 + 0.5/50) = P(Z \leq (0.02 + 0.5/50 - 0.05) / 0.031) - P(Z \leq (0.00 - 0.5/50 - 0.05) / 0.031) = P(Z \leq -0.64) - P(Z \leq -1.93) = 23.43\%$

Observe que los valores $\pm$ 0.5 / 50, constituyen los valores de $\pm$ 0.5, para ajustar los valores discretos de la distribución binomial a valores continuos propios de la distribución normal, sólo que se han transformado en términos de fracción defectuosa al dividirlos por n, que en este caso es 50.

Ejemplo 5: Un lote de 50 artículos tiene 5 defectuosos y 45 no defectuosos. Se toma una muestra de 10 artículos, cuál es la probabilidad de obtener tres o más defectuosos?

Objetivo del ejemplo: Mostrar la relación entre la proporción poblacional y la proporción de muestras usando el factor de corrección y su aplicación en el calculo de probabilidades.

Solución:

$\mu_p = p = 0.10$ y $\sigma_p = \sqrt{\frac{pq}{n} x \frac{N-n}{N-1}} = 0.0857$

La probabilidad de que obtener 3 o más defectuosos es equivalente a obtener entre 3 y 10 defectuosos que corresponden a los valores de P de 0.3 y 1.0, por lo tanto, la probabilidad pedida, es: P(0.3 - 0.5 /10 ≤ P ≤ 1.0 + 0.5 /10) = P(0.25 – 0.1 / 0.0857 ≤ Z ≤ 1.05 – 0.1 / 0.0857) = P(Z ≤11.08) – P(Z ≤ 1.75) = 1- 0.9599 = 4.01%

6. ESTIMACIÓN DE PARÁMETROS

La aplicación de la teoría de muestreo, usada en los numerales anteriores, asumía que los parámetros poblacionales eran conocidos y a partir de éstos se calculaban los valores de los estadísticos de muestras.

En muchas ocasiones el valor de los parámetros no se conoce y la teoría del muestreo puede usarse para estimarlos a partir de la observación de muestras extraídas de la población.

La estimación de parámetros a partir de estadísticos de muestras puede hacerse bajo dos modalidades: Estimación puntual y estimación de intervalo. La estimación puntual consiste en la determinación de un valor único del parámetro. En la estimación de intervalo se determina un rango dentro del cual se espera esté el valor del parámetro.

La estimación de parámetros es un campo muy amplio en la teoría estadística y escapa al alcance de este texto (véase: Meyer Paul. Probabilidad y Aplicaciones Estadísticas, capitulo 14). En este texto sólo se tratará el caso de estimación puntual para los parámetros media, desviación estándar y proporciones.

- **Estimación de la media poblacional μ:** La media poblacional puede estimarse a partir de la media de una muestra de tamaño $n \geq 30$ o alternativamente al promedio de los promedios de k muestras de tamaño n, en este caso, k deberá ser mayor o igual que 20 y el valor de n en términos generales no tiene restricciones.

En este último caso se sigue el siguiente procedimiento:

- Se extraen k muestras de tamaño n
- Se calcula el promedio de cada una de las k muestras

$$\xi_K = \frac{\sum_{i=1}^{n} x_i}{n}$$

- Se calcula el promedio de los k promedios anteriores

$$\overline{\overline{X}} = \frac{\sum_{J=1}^{K} x_J}{k}$$

- $\overline{\overline{X}}$ es un estimador de μ.

- **Estimación de la varianza poblacional σ^2 :** Para la estimación de la varianza poblacional se sigue el siguiente procedimiento:

- Se toman k muestras de n elementos cada una
- Se calcula la varianza, s^2, de cada una de la k muestras

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1} \qquad \text{o} \qquad s^2 = \frac{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)}$$

- Se calcula el promedio de las k varianzas de las muestras

$$\overline{S}^2 = \frac{\sum_{K=1}^{K} S^2{}_K}{k}$$

- $\overline{S}^2$ es un estimador de σ^2

Ejemplo 6: Una máquina produce un tipo de rodamiento cuyo peso promedio y desviación varianza son desconocidos. Se decide tomar muestras de 5 rodamientos durante 20 días para estimar el peso promedio y la varianza de los rodamientos producidos por la máquina.

Objetivo del ejemplo: Ilustrar los procedimientos para la estimación de la media y la varianza poblaciones a partir de la observación de muestras.

La siguiente tabla muestra los valores obtenidos en las 20 muestras (en Kg) y los cálculos de las medias y varianzas de las muestras.

DIA	OBS.1	OBS.2	OBS.3	OBS.4	OBS.5	MEDIA	VAR
1	0,75	0,76	0,77	0,75	0,74	0,754	0,00013
2	0,77	0,75	0,77	0,81	0,76	0,772	0,00052
3	0,74	0,74	0,79	0,74	0,79	0,760	0,00075
4	0,78	0,78	0,80	0,72	0,73	0,762	0,00122
5	0,74	0,79	0,75	0,74	0,73	0,750	0,00055
6	0,74	0,73	0,76	0,74	0,77	0,748	0,00027
7	0,75	0,78	0,77	0,72	0,77	0,758	0,00057
8	0,77	0,79	0,72	0,74	0,79	0,762	0,00097
9	0,74	0,73	0,75	0,78	0,80	0,760	0,00085
10	0,72	0,76	0,77	0,79	0,75	0,758	0,00067
11	0,74	0,74	0,79	0,73	0,74	0,748	0,00057
12	0,72	0,78	0,72	0,78	0,78	0,756	0,00108
13	0,80	0,79	0,74	0,79	0,79	0,782	0,00057
14	0,75	0,73	0,74	0,80	0,73	0,750	0,00085
15	0,80	0,79	0,80	0,72	0,78	0,778	0,00112
16	0,76	0,73	0,75	0,74	0,79	0,754	0,00053
17	0,79	0,76	0,76	0,74	0,73	0,756	0,00053
18	0,73	0,74	0,73	0,78	0,75	0,746	0,00043
19	0,81	0,75	0,76	0,79	0,76	0,774	0,00063
20	0,77	0,79	0,74	0,73	0,73	0,752	0,00072
			PROMEDIOS			0,759	0,00068

Tabla6.4. Datos y cálculos del ejemplo 6

Solución:

Dado que $\overline{\overline{X}}$ = 0.759 y $\overline{S}^2$ = 0.00068, estos valores son estimadores de μ y σ^2 de la producción de la máquina respectivamente.

- **Estimación de la proporción poblacional P:** La estimación de la proporción poblacional puede hacerse a partir del estadístico p, tomando una muestra representativa de tamaño n.

Ejemplo 7: Se desea estimar la proporción de electores que votarán a favor de un candidato en las próximas elecciones municipales. Para ello se consulta a 5000 electores, de los cuales 1780 manifiestan favoritismo por el candidato.

Objetivo del ejemplo: Mostrar el proceso de la estimación de la proporción poblacional a partir de la información obtenida en una muestra.

Solución: p = 1780 / 5000 = 35.6%. Este valor es un estimador de la proporción poblacional P.

De manera alternativa, la estimación de la proporción poblacional, puede hacerse tomando k muestras de tamaño n y calculando el promedio de las k proporciones de las muestras, $\overline{p} = \dfrac{\sum_{K=1}^{K} p_K}{k}$ y el valor $\overline{p}$ es un estimador de P.

Ejemplo 8: Un proceso continuo produce un artículo que puede clasificarse como bueno o defectuoso.

Con el fin de estimar la proporción de defectuosos del proceso, se toman muestras de 50 unidades durante 20 días. Los resultados del muestreo se presentan en la tabla:

MUESTRA Nº	NUMERO DEFECTUOSOS	FRACCIÓN DEFECTUOSA p
1	2	0,04
2	1	0,02
3	2	0,04
4	1	0,02
5	0	0,00
6	2	0,04
7	3	0,06
8	0	0,00
9	2	0,04
10	0	0,00
11	3	0,06
12	1	0,02
13	1	0,02
14	2	0,04
15	3	0,06
16	1	0,02
17	2	0,04
18	2	0,04
19	0	0,00
20	2	0,04
N = 50	PROMEDIO	0,030

Tabla 6.5. - Datos del ejemplo 8

Objetivo del ejemplo: Ilustrar la estimación de la proporción de defectuosos de un proceso productivo a partir de la extracción de muestras.

Solución: De la tabla puede observarse que $\bar{p}$ = 3% y este valor es un estimador de P del proceso.

7. TALLER

1. En una caja hay cinco balotas marcadas con los números: 0, 2, 4, 6 y 8. Si se define la variable aleatoria X como el valor marcado en las balotas.

- Calcule la media y la desviación estándar de la población.
- Determine la distribución de probabilidad muestral de la media al extraer 2 balotas al azar con sustitución.
- Calcule la media y la varianza de la anterior distribución.
- Determine la relación entre las medias y las desviaciones estándar poblacionales y de las muestras.

2. En relación con el problema anterior

- Determine la distribución de probabilidad de muestras de la media al extraer 2 balotas al azar sin sustitución.
- Calcule la media y la varianza de la anterior distribución y relaciónelas con los parámetros.

3. La estatura de los estudiantes de la Facultad se distribuye aproximadamente normal con media de 1.70 m y desviación estándar de 0.15 m. Calcule las siguientes probabilidades:

- De que la estatura promedio de 100 estudiantes elegidos al azar sea mayor o igual a 1.73 m.
- De que la estatura promedio de 50 estudiantes elegidos al azar sea mayor o igual a 1.73 m.
- De que la estatura promedio de 25 estudiantes elegidos al azar sea mayor o igual a 1.73 m.
- De que la estatura promedio de 1 estudiante elegido al azar sea mayor o igual a 1.73 m.

4. Un granjero tiene un hato de 2000 novillos. El peso de los novillos se distribuye normal con media de 350 kg y desviación estándar de 25 kg. Un comprador desea comprarle algunos novillos pero le advierte que los animales que le despache estén entre los 400 novillos más pesados del hato. Al día siguiente le solicitó 16 novillos. El granjero ignorando la advertencia selecciono del hato 16 novillos al azar y los despachó.

- Cuál es la probabilidad de que los novillos despachados cumplan la condición impuesta por el comprador?

- Cuál es la probabilidad de que los novillos despachados cumplan la condición impuesta por el comprador, si el pedido es de 4 novillos en vez de 16.

5. Un grupo de estudiantes de estadística está realizando una práctica sobre distribuciones de muestras. Algunos de los resultados logrados son:

El número de abuelos vivos de los trabajadores del puerto se distribuye aproximadamente normal.

Cuatro de los trabajadores no tienen abuelos vivos.

La distribución de muestras del promedio de abuelos vivos para muestras de dos trabajadores, es:

Promedio	0	0.5	1	1.5	2	2.5	3	3.5	4
Probabilidad	0.01	0.02	0.08	0.21	0.36	0.18	0.08	0.05	0.01

- ¿Cuántos trabajadores tiene el puerto?

- ¿Cuántas muestras de dos trabajadores tienen promedio de 1.5 o menos abuelos vivos?

6. Una máquina llena sacos de harina cuyo peso se distribuye normal con media, μ = 50 kg y desviación estándar, σ = 1.75 kg. Si se toma una muestra de 3 sacos, cuáles son las probabilidades de que el promedio de la muestra caiga en los rangos:

- $[\mu \pm 3\sigma]$
- $[\mu \pm 2\sigma]$
- $[\mu \pm 1.5\sigma]$
- $[\mu \pm 0.5\sigma]$

7. Si en el problema anterior un saco se considera defectuoso si pesa más 51 kilogramos o menos de 49.5 kilogramos, cuál es la probabilidad de que 1 saco tomado al azar sea defectuoso?

8. Un distribuidor de cremalleras tiene un lote 2.5% defectuoso de 3000 unidades. Recibe un pedido de un cliente. El cliente rechaza el pedido si al examinar 50 cremalleras encuentra una fracción defectuosa igual o mayor al 4%. Cuál es la probabilidad de que le rechacen el lote?

9. En relación con el problema anterior, otro cliente rechaza el lote, si al examinar 75 cremalleras encuentra 3 o más defectuosas. Cuál es la probabilidad de que le sea rechazado el lote?

10. Se tiene un lote de un millón de botones 2% defectuoso

- Cuál es la probabilidad de que 1000 de estos botones extraídos al azar tengan una fracción defectuosa entre 1.5% y 2.5%.
- Cuál es la probabilidad de que 1000 de estos botones extraídos al azar tengan una fracción defectuosa menor o igual al 1%.
- Cuántos botones deberán extraerse del lote para que la probabilidad de que la muestra tenga una fracción defectuosa menor o igual al 1% sea del 10%

8. BIBLIOGRAFÍA

- Meyer, P. *Probabilidad y Aplicaciones Estadísticas.* Fondo Educativo Interamericano. México 1981.
- Ash, R. *Basic Probability Theory.* John Wiley & Sons. New York. 1976.
- Walpole, R. *Introduction to Statistics.* Macmillan Company.New York 1969.
- Mason, R. y D. Lind. *Estadística para Administración y Economía.* Alfaomega. México 1992
- Mood, A. y F. Graybill. *Introducción a la Teoría de la Estadística.* Aguilar. Madrid. 1972

SOLUCION A LOS PROBLEMAS

TALLER MODULO 1 ESPACIOS DE MUESTRAS Y SUCESOS

Problema 1:

- Universal: U = {Luis, Dora, Rene, Lina, Inés, Rita, Pepe, Roque, José, Guido}
- Tienen carro: C = {Luis, Dora, Inés, Pepe, Roque}
- Tienen moto: M = {Luis, Rene, Rita, Dora}
- No tiene Carro: NC = {Rene, Lina, Rita, José, Guido}
- No tienen moto: NM = {Lina, Inés, Pepe, Roque, José, Guido}
- Tienen carro y moto: CyM = $C \cap M$ = {Luis, Dora}
- No tiene carro ni moto: NCyNM = $NC \cap NM$ = {Lina, José, Guido}
- Tienen carro o tienen moto: CoM = $C \cup M$ = {Luis, Dora, Rene, Inés, Rita, Pepe, Roque}
- Tienen carro y no tienen moto: CyNM = $C \cap NM$ = {Inés, Pepe, Roque}
- Tienen moto y no tienen carro: $M \cap NC$ = {Rene, Rita}
- No tienen moto o no tienen carro: NMoNC = $NM \cup NC$ = {Rene, Lina, Inés, Rita, Pepe, Roque, José, Guido}
- Tienen carro o no tienen ni moto ni carro: $C \cup (NM \cap NC)$ = {Luis, Dora, Lina, Inés, Pepe, Roque, José, Guido}

Problema 2: Cada dígito de la placa puede ser de 10 maneras, por la regla de multiplicación, el número total de placas es: $10 \times 10 \times 10 \times 10 \times 10 = 10^5 = 100000$.

Problema 3: La primera posición de la placa puede ser de 5 maneras, la segunda de 5 maneras, la tercera de 10 maneras, la cuarta de 10 maneras y la quinta de 10 maneras, por lo tanto, el total de placas, es: $5 \times 5 \times 10 \times 10 \times 10 = 25000$

Problema 4:

- Las camas pueden ser ocupadas de ${}_{10}P_2$ = 90 maneras

- La habitación puede ser ocupada de ${}_{10}C_2$ = 45 maneras

- Para que puedan comunicarse deberán ocupar la habitación dos de los tres que hablan chino o dos de los siete que hablan español. Esto es, ${}_3C_2 + {}_7C_2$, es decir, 3 + 21 = 24 maneras.

Problema 5:

- La primera choza puede ser ocupada de 8 maneras, la segunda de 7 maneras, la tercera de 6 maneras y así sucesivamente, por la regla de multiplicación, el total de maneras es: 8x7x6x......x1 = 8! = 40320 maneras.

- Para 4 chozas se tiene: La primera choza puede ser ocupada de ${}_8C_2$ = 28 maneras, por cada una de estas maneras la segunda choza puede ser ocupada de ${}_6C_2$ = 15 maneras, por cada una de las anteriores ocupaciones la tercera choza puede ser ocupada de ${}_4C_2$ = 6 maneras y finalmente por cada una de las anteriores ocupaciones la cuarta choza puede ser ocupada de ${}_2C_2$ = 1 manera. Por lo tanto el total de maneras es: 28x15x6x1 = 2520 maneras.

- Para el caso de 2 chozas de 5 y 3 ocupantes, se tiene: El número de grupos de 5 personas entre los ocho náufragos es: ${}_8C_5$ = 56 y por cada uno de estos grupos, los tres restantes “agrupados” de a tres es ${}_3C_3$ = 1, por lo tanto, las chozas pueden ser ocupadas de 56x1 = 56 maneras.

Problema 6: El total de mensajes posibles es: ${}_5P_2$ = 20 mensajes

Problema 7:

- Con 5 banderas y código de 3 banderas: ${}_5P_3$ = 60 mensajes

- Con 6 banderas y código de 2 banderas: ${}_6P_2$ = 30 mensajes

- Con 6 banderas y código de 3 banderas: ${}_6P_3$ = 120 mensajes

Problema 8:

- Sean los sucesos:

$V_{i,J}$ = No Muere (queda vivo) el jugador i en la ronda j.
$M_{i,J}$ = Muere el jugador i en la ronda j.

El espacio muestral S = $\{(V_{1-1}, V_{2-1}, V_{1-2}, V_{2-2}), (V_{1-1}, V_{2-1}, V_{1-2}, M_{2-2}), (V_{1-1}, V_{2-1}, M_{1-2}), (V_{1-1}, M_{2-1}), (M_{1-1})\}$

- Igual que el punto anterior.

Problema 9:

- ${}_{50}C_4$ = 230300 maneras
- ${}_{18}C_2$ = 153 maneras
- ${}_{32}C_2$ = 496 maneras
- ${}_{18}C_2 \times {}_{32}C_2$ = 153 x 496 = 75888 maneras

Problema 10: ${}_{42}C_6$ = 5245786 arreglos diferentes

Problema 11: La primera rueda puede caer de 10 maneras, la segunda de 10 maneras, la tercera de 10 maneras, la cuarta de 12 maneras y la quinta de 5 maneras. Por lo tanto, el total de billetes diferentes es: 10x10x10x12x5 = 60000.

Problema 12:

- En la primera ronda hay 8 partidos y por lo tanto 8 ganadores. El número total de maneras de 8 ganadores entre los 16 jugadores, es: ${}_{16}C_8$ = 12870
- De los 16, hay 14 extranjeros, de los cuales habrá 6 ganadores (8 – 2), por lo tanto, ${}_{14}C_6$ = 3003

TALLER MODULO 2 PROBABILIDAD

Problema 1: Usando la frecuencia relativa se tiene: P(llueva) = 73 / 500 = 14.6%

Problema 2:

- P(3 defectuosos) = 0.18 x 0.18 x 0.18 = 0.583% (Sucesos independientes)
- P(1º def) x P(2º bueno) x P(3º bueno) = 0.18 x 0.82 x 0.82 = 12.1%
- P(2 de 3 def) = P[(1º def $\cap$ 2º def $\cap$ 3º bueno) $\cup$ (1º def $\cap$ 2º bueno $\cap$ 3º def) $\cup$ (1º bueno $\cap$ 2º def $\cap$ 3º def)] = (0.18x0.18x0.82) + (0.18x0.82x018) + (0.82x0.18x018) = 7.97%

Problema 3: Número de casos totales: ${}_5P_3$ = 60. Número de casos favorables: ${}_3P_3$ = 6. Por lo tanto, la probabilidad pedida es: 6 / 60 = 10%

Problema 4: Sean los sucesos:

Perno del proveedor 1 → p1
Perno del proveedor 2 → p2
Perno del proveedor 3 → p3
Perno defectuoso → pd

Se conocen: P(p1) = 1000 / 2300 = 43.48%
P(p2) = 500 / 2300 = 21.74%
P(p3) = 800 / 2300 = 34.78%
P(pd) = (150 + 60 + 64) / 2300 = 11.91%
P(pd / p1) = 15%
P(pd / p2) = 12%
P(pd / p3) = 8%

Se pide calcular: P(p1 / pd); P(p2 / pd) y P(p3 / pd)

Se pueden calcular: P(pd $\cap$ p1) = P(pd/p1) x P(p1) = 6.52%
P(pd $\cap$ p2) = P(pd/p2) x P(p2) = 2.61%
P(pd $\cap$ p3) = P(pd/p3) x P(p3) = 2.78%

Por lo tanto: Usando la regla de Bayes, se tiene:

P(p1/pd) = 0.0652 / (0.0652 + 0.0261 + 0.0278) = 54.74%
P(p2/pd) = 0.0261 / (0.0652 + 0.0261 + 0.0278) = 21.92%
P(p3/pd) = 0.0278 / (0.0652 + 0.0261 + 0.0278) = 23.34%

La probabilidad de que el perno provenga del proveedor 1 es la mayor.

Problema 5: Sean los sucesos: Caja 1 → c1; Caja 2 → c2 y Caja3 → c3
Bola blanca → b; Bola negra → n

Se conocen: P(c1) = 1 / 3; P(c2) = 1 / 3; P(c3) = 1 / 3
P(b/c1) = 1; P(b/c2) = 0; P(b/c3) = 1 / 2

Se pide: P(c1/b)

Puede calcularse: $P(b \cap c1) = 1 \times 0.3333 = 0.3333$
$P(b \cap c2) = 0$
$P(b \cap c3) = 0.5 \times 0.333 = 0.1665$

Luego, P(c1/b) = 0.333 / (0.333 + 0.165) = 66.66%

Este valor corresponde a que la segunda bola sea blanca.

Problema 6: Sean los sucesos:

$V_{i,J}$ = No Muere (queda vivo) el jugador i en la ronda j.
$M_{i,J}$ = Muere el jugador i en la ronda j.

El espacio muestral $S = \{(V_{1-1}, V_{2-1}, V_{1-2}, V_{2-2}), (V_{1-1}, V_{2-1}, V_{1-2}, M_{2-2}), (V_{1-1}, V_{2-1}, M_{1-2}), (V_{1-1}, M_{2-1}), (M_{1-1})\}$

- El suceso de interés: Que ambos queden ilesos y girando el tambor, es: $(V_{1-1}, V_{2-1}, V_{1-2}, V_{2-2})$.

  Y su probabilidad asociada , es: $P(V_{1-1} \cap V_{2-1} \cap V_{1-2} \cap V_{2-2}) = (5/6)^4 = 48.22\%$

- El suceso de interés: Que ambos queden ilesos sin girar el tambor, es: $(V_{1-1}, V_{2-1}, V_{1-2}, V_{2-2})$.

  Y su probabilidad asociada , es: $P(V_{1-1} \cap V_{2-1} \cap V_{1-2} \cap V_{2-2}) = 5/6 \times 4/5 \times 3/4 \times 2/3 = 33.33\%$

Problema 7: Se tiene: P(llueva) = 0.40; P(pesca buena / llovió) = 0.90.
Se pide: P(pesca buena $\cap$ llovió) = 0.90 x 0.40 = 36%

Problema 8: Se tiene: P(pesca buena / llovió) = 0.75%
P(llovió / pesca buena) = 0.65%
P(llovió $\cap$ pesca buena) = 0.55%

- El régimen de lluvias es: P(llueva) = 0.55 / 0.75 = 73.33%

- Se pide: P(buena pesca) = 0.55 / 0.65 = 84.6%, por lo tanto en los próximos 100 días se esperarían 85 días de buena pesca.

Problema 9: Si acepta la segunda elección, se tiene:

Baúl 1: Cheques de \$ 200000 y \$ 50000
Baúl 2: Cheques de \$ 1800000 y \$ 450000

P(Baúl 1) = 0.50; P(Baúl 2) = 0.50

P(\$ 200000) = P(Baul 1 $\cap$ \$ 200000) = 0.5 x 0.5 = 0.25
P(\$ 50000) = P(Baul 1 $\cap$ \$ 50000) = 0.5 x 0.5 = 0.25
P(\$1800000) = P(Baul 2 $\cap$ \$1800000) = 0.5 x 0.5 = 0.25
P(\$ 450000) = P(Baul 2 $\cap$ \$ 450000) = 0.5 x 0.5 = 0.25

El valor esperado de la segunda elección es:

(200000x0.25) + (50000x0.25) + (1800000x0.25) + (450000x0.25)= 625000.

Valor mayor a los \$ 600000 seguros, por lo tanto debe aceptar la segunda elección.

Problema 10: Sean: C → Caspirina; S → Saspirina y R → Raspirina.

Se tiene: P(C) = 20 / 115 = 17.40% P(sane/C) = 75%
P(S) = 50 / 115 = 43.47% P(sane/S) = 45%
P(R) = 45 / 115 = 39.13% P(sene/R) = 65%

Se pide hallar: P(C/sanó) ; P(S/sanó) y P(R/sanó)

Pueden calcularse: P(sanó $\cap$ C) = 0.75 x 0.1740 = 13.05%
P(sanó $\cap$ S) = 0.45 x 0.4347 = 19.56%
P(sanó $\cap$ R) = 0.65 x 0.3913 = 25.43%

Las probabilidades pedidas son:

P(C/sanó) = 0.1305 / (0.1305 + 0.1956 + 0.2543) = 22.48%
P(S/sanó) = 0.1956 / (0.1305 + 0.1956 + 0.2543) = 33.70%
P(R/sanó) = 0.2543 / (0.1305 + 0.1956 + 0.2543) = 43.82%

La probabilidad de haber tomado raspirina es la mayor.

Problema 11:

- P(acertar) = 1 / ${}_{42}C_6$ = 1 / 5245786
- Casos favorables: ${}_{21}C_6$ = 54264, por lo tanto: P(todos pares) = 54264 / 5245786 = 1.03%

Problema 12:

- Casos totales: ${}_{5}P_5$ = 120. P(orden alfabético ascendente) = 1/120 = 0.83%
- P(orden alfabético descendente) = 1/120 = 0.83%
- P(ascendente $\cup$ descendente) = 0.833 + 0.833 = 1.66%

TALLER MODULO 3 VARIABLE ALEATORIA Y DISTRIBUCIONES DE PROBABILIDAD

Problema 1:

- X = {1, 2, 3, 4, 5, 6}
- X = {2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}
- X = {1, 2, 3, 4, 5, 6, 8, 9, 10, 12, 15, 16, 18, 20, 24, 25, 30, 36}
- X = {0, 1, 2, 3, 4, 5}
- X = {0, 1, 2, 3}
- X = {1, 2, 3, 4, 5, 6, 7}

Problema 2:

V. A. (x)	1	2	3	4	5	6
P(X = x)	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

V. A. (x)	2	3	4	5	6	7	8	9	10	11	12
P(X = x)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

VA (x)	1	2	3	4	5	6	8	9	10	12	15	16	18	20	24	25	30	36
P(X=x)	1/36	2/36	2/36	3/36	2/36	4/36	2/36	1/36	2/36	4/36	2/36	1/36	2/36	2/36	2/36	1/36	2/36	1/36

V. A. (x)	0	1	2	3	4	5
P(X = x)	6/36	10/36	8/36	6/36	4/36	2/36

V. A. (x)	0	1	2	3
P(X = x)	1/8	3/8	3/8	1/8

V. A. (x)	1	2	3	4	5	6	7
P(X = x)	1/12	2/12	2/12	2/12	2/12	2/12	1/12

Problema 3:

V. A. (x)	-3	-2	-1	1	2	4	5	9
P(X = x)	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8

V. A. (x)	1	4	9	25	81
P(X = x)	2/7	2/7	1/7	1/7	1/7

V. A. (x)	1	2	3	4	5	9
P(X = x)	2/8	2/8	1/8	1/8	1/8	1/8

Problema 4:

- $\sum P(x)$ = 1/10 + 1/10 + 3/10 + 2/10 + 1/10 + 2/10 = 1

- E(x) = 0/10 + 7/10 + 24/10 + 22/10 + 12/10 + 28/10 = 9.3

  V(x) = (49/10 + 192/10 + 242/10 + 144/10 + 392/10) – 9.3^2 = 15.41

Problema 5:

- X = {1, 2, 3, 4}

- P(X = 1) = 2/5 = 0.4
  P(X = 2) = 3/5 x 2/4 = 6/20 = 0.3
  P(X = 3) = 3/5 x 2/4 x 2/3 = 12/60 = 0.2
  P(X = 4) = 3/5 x 2/4 x 1/3 x 2/2 = 12/120 = 0.1

V. A. (x)	1	2	3	4
P(X = x)	0.4	0.3	0.2	0.1

- E(X) = 1x0.4 + 2x0.3 + 3x0.2 + 4x0.1 = 2
  V(X) = (1x0.4 + 4x0.3 + 9x0.2 + 16x0.1) – 2^2 = 5 – 4 = 1

Problema 6: P(ganar) = 1/10000; P(perder) = 9999/10000
Si gana: \$ 1000000 – \$ 500 = \$ 999500. Si pierde: \$ - 500

Valor esperado = 999500/10000 – 500x9999/10000 = - \$ - 400

Problema 7: El espacio muestral (G → germina; M → no germina), es:

S = {MMM, MMG, MGM, GMM, MGG, GMG, GGM, GGG}

- P(0 germinen) = 0.3^3 = 0.027
  P(1 germine) = 0.7 x 0.3^2 x 3 = 0.189
  P(2 germinen) = 0.7^2 x 0.3 x 3 = 0.441
  P(3 germinen) = 0.7^3 = 0.343

V. A. (x)	0	1	2	3
P(X = x)	0.0127	0.189	0.441	0.343

- E(X) = 1x0.189 + 2x0.441 + 3x0.343 = 2.1

  V(X) = (1x0.189 + 4x0.441 + 9x0.343) – 2.1^2 = 0.63; σ = 0.7937

Problema 8:

-

V. A. (x)	2	3	6
P(X = x)	1/2	1/3	1/6

- $\sum \mathrm{P(x)}$ = 1/2 + 1/3 + 1/6 = 1

- E(X) = 2/2 + 3/3 + 6/6 = 3
  V(X) = (4/2 + 9/3 + 36/6) – 3^2 = 2

TALLER MODULO 4 DISTRIBUCIONES DE PROBABILIDAD DISCRETAS

Problema 1:

- Marca A: En este caso el "experimento" es sacar un grano y verificar si "estalla o no". La probabilidad de que estalle es del 99% y esta probabilidad permanece constante en todas las repeticiones del experimento, es decir es un experimento binomial.

  Marca B: En este caso de los 100 granos hay 99 que estallan y al repetir el experimento en 99 de estas repeticiones la probabilidad de que estalle es 1 y en una de las repeticiones la probabilidad de que estalle es cero.

- Marca A: La distribución de probabilidad es binomial, con n =100 y p = 1%, por lo tanto, la probabilidad de que no estallen 0, 1, 2, ..., k, está dada por b(k, 100; 0.01)

  Marca B: La distribución de probabilidad es hipergeométrica, h(1; 100; 100; k), observe que la probabilidad de que X granos no estallen es 1 cuando X = 1 y 0 cuando $X \neq 1$.

- Marca A: $\mu = np = 100x\ 0.01 = 1$

  Marca B: $\mu = nk / N = 100/100 = 1$

- P(0 no estallen) = 0; P(1 no estalle) = 1;
  P(2 no estallen) = 0; P(3 no estallen) = 0.

- P(0 no estallen) = h(0;100; 50; 1) = 0.5
  P(1 no estalle) = h(1; 100; 50; 1) = 0.5
  P(2 no estallen) = 0
  P(3 no estallen) = 0

Problema 3: h(3; 10; 3; 3) = 0.83%

Problema 4:

- h(0: 5; 2; 3) + h(1; 5; 2; 3) = 0.1 + 0.6 = 70%

- h(0: 100; 10; 3) + h(1; 100; 10; 3) = 0.7265 + 0.2476 = 97.41%

- h(0: 100; 10; 30) + h(1; 100; 10; 30) = 0.0229 + 0.1127 = 13.56

En los tres casos se usó la distribución hipergeométrica dado que n/N >5%

Problema 5:

- P(0 pares) = b(10; 10; 0.6) + b(0; 10; 0.6) = 0.0060 + 0.0001 = 0.61%
- P(1 par) = b(1; 10; 0.6) + b(9; 10; 0.6) = 0.0016 + 0.0403 = 4.19%
- P(2 pares) = b(2; 10; 0.6) + b(8; 10; 0.6) = 0.0106 + 0.1209 = 13.15%
- P(3 pares) = b(3; 10; 0.6) + b(7; 10; 0.6) = 0.0425 + 0.2150 = 25.75%
- P(4 pares) = b(4; 10; 0.6) + b(6; 10; 0.6) = 0.1115 + 0.2508 = 36.23%
- P(5 pares) = b(5; 10; 0.6) = 0.2006 = 20.06%

Problema 6:

- Experimento: Determinar si la respuesta es buena (b), o mala (m).

  Repeticiones: 10

  P(b) = 0.25; P(m) = 0.75. Permanecen constantes en las 10 repeticiones

  Espacio muestral: Respuestas buenas, S = {0, 1, 2, ..., 10}

  Suceso: Sacar 6 o más buenas para no salir. A = {5, 6, 7, 8, 9, 10}

- Distribución Binomial.

- P(A) = $\sum_{k=5}^{10} b(k,10,0.25)$ = 7.81%

Problema 7: Distribución Poisson, con α = 2. k = 4

P(X = 4) = 9.02%

Problema 8:

- P(sacar 1) = 1/6. Permanece constante en las n = 3 repeticiones.

  P(sacar tres 1) = b(3; 3; 1/6) = 0.46%

- Usando Poisson: α = 3 x 1/6 = 0.5. k = 3. P(tres 1) = 1.26%

- Los resultados son muy diferentes, la aproximación usando la distribución Poisson no es válida, puesto que la muestra es muy pequeña y p muy grande.

Problema 9:

- Con p = 20%. Usando la distribución binomial:

  P(9 o más buenos) = b(0; 10; 0.2) + b(1; 10; 0.2) = 37.58%

  Usando la distribución Poisson:

  P(9 o más buenos) = 40.60%

- Con p = 30%. Usando la distribución binomial:

  P(9 o más buenos) = b(0; 10; 0.3) + b(1; 10; 0.3) = 14.93%

  Usando la distribución Poisson:

  P(9 o más buenos) = 19.91%

- Con p = 10%. Usando la distribución binomial:

  P(9 o más buenos) = b(0; 10; 0.1) + b(1; 10; 0.1) = 73.61%

  Usando la distribución Poisson:

  P(9 o más buenos) = 73.57%

- A medida que la fracción defectuosa disminuye la aproximación entre las distribuciones mejora.

TALLER MODULO 5 DISTRIBUCION NORMAL

Problema 1: Se tiene: $N = 3000$; $\mu = 3.2$; $\sigma = 0.4$

- $P(X \geq 4) = 1 - P(Z < (4 – 3.2) / 0.4) = 1 – P(Z < 2) = 1 – 0.9772 = 2.28\%$ y el número de estudiantes: $2.28\% \times 3000 \approx 68$

- $P(X \leq 2) = P(Z \leq (2 – 3.2) / 0.4) = P(Z \leq -3) = 0.13\%$, y el número de estudiantes: $0.13\% \times 3000 \approx 4$

- $P(2.5 \leq X \leq 3.5) = P(Z < (3.5 – 3.2) / 0.4) - P(Z < (2.5 – 3.2) / 0.4) = P(Z < 0.75) - P(Z < - 1.75) = 0.7734 – 0.0401 = 73.33\%$, y el número de estudiantes: $73.33\% \times 3000 \approx 2200$

- $P(Z \geq (x – 3.2) / 0.4) = 150/3000 = 0.05 \rightarrow (x – 3.2) / 0.4) = 1.645$, por lo tanto, $x \approx 3.86$

- $P(Z \geq (3.5 – 3.2) / \sigma) = 0.25 \rightarrow 0.3 / \sigma \approx 0.67 \rightarrow \sigma \approx 0.448$

Problema 2: Se tiene: $\mu = 170$; $\sigma = 10$

- $1 - P(169 \leq X \leq 173) =1 - P((169 – 170) / 10 \leq Z \leq (173 - 170) / 10) = 1 - P(-0.1 \leq Z \leq 0.3) = 1 – (0.6179 – 0.4602) = 84.23\%$

- $0.8423 \times (1- 0.8423) = 13.28\%$

Problema 3: Se tiene: $\mu = 20$; $\sigma = 2$

- $P(X \leq 18) = P(Z \leq (18 – 20) / 2) = P(Z \leq - 1) = 15.87\%$

  $P(18 \leq X \leq 20) = P(Z < (20 – 20) / 2) - P(Z < (18 – 20) / 2) = P(Z < 0) - P(Z < - 1) = 0.5000 – 0.1587 = 34.13\%$

  $P(X \geq 20) = 1 - P(Z < (20 – 20) / 2) = 1 – P(Z < 0) = 1 – 0.5000 = 50\%$

  El valor esperado: $0 \times 15.87\% + 1500 \times 34.13\% + 2000 \times 50\% \approx \$ 1512$

Problema 4: Se tiene: $\mu = 1000$; $P(X \geq 942) = 98\%$

$1 - P(Z \geq (942 – 1000) / \sigma) = 2\% \rightarrow - 8 / \sigma = - 0.85$, por lo tanto, $\sigma \approx 9.41$

Problema 5: Se tiene: $\mu = 35$; $\sigma = 4$

$P(X \geq 40) = 1 - P(Z < (40 – 35) / 4) = 1 – P(Z < 1.25) = 10.56\%$

Problema 6: Se tiene: $\mu_A = 750$; $\sigma_A = 45$ y $\mu_B = 780$; $\sigma_B = 18$

Para batería A: $P(X \geq 800) = 1 - P(Z < (800 – 750) / 45) = 1 – P(Z < 1.11) = 1 – 0.8665 = 13.35\%$

Para batería B: $P(X \geq 800) = 1 - P(Z < (800 – 780) / 18) = 1 – P(Z < 1.11) = 1 – 0.8665 = 13.35\%$

Es indiferente comprar cualquiera de las dos marcas.

Problema 7: Se tiene: $\mu_A = 18$; $\sigma_A = 1.5$ y $\mu_B = 18$

- $P(X \leq 17) = P(Z \leq (17 – 18) / 1.5) = P(Z \leq - 0.66) = 25.46\%$
- $P(X \leq 17) = P(Z \leq (17 – 18) / \sigma_B) = 12.73\% \rightarrow -1 / \sigma_B = -1.14$; $\sigma_B \approx 0.877$

Problema 8: Se tiene: $\mu = 1.74$ y $\sigma = 0.12$

- $P(1.74–0.12 \leq X \leq 1.74+0.12) = P(Z<(1.86–1.74)/0.12) - P(Z<(1.66–1.74)/0.12) = P(Z < 1) - P(Z < -1) = 0.8413 – 0.1587 = 68.26\%$
- $P(Z < 2) - P(Z < -2) = 0.9772 – 0.0228 = 95.44\%$
- $P(Z < 3) - P(Z < -3) = 0.9987 – 0.0013 = 99.74\%$

Problema 9: Se tiene: $\mu = 42$ y $\sigma = 2.5$

- $P(X \leq 38) = P(Z \leq (38 - 42) / 2.5) = P(Z \leq -1.6) = 5.48\%$

TALLER MODULO 6 TEORIA DEL MUESTREO

Problema 1:

- $\mu = (0 + 2 + 4 + 6 + 8) / 5 = 4$

  $\sigma^2 = (16 + 4 + 0 + 4 + 16) / 5 = 8 \rightarrow \sigma = 2.8284$

- Usando la regla de multiplicación el número de elementos del espacio muestral es: 5 x 5 = 25. En la siguiente tabla se muestran las medias, ξ, de estas 25 duplas:

	0	2	4	6	8
0	0	1	2	3	4
2	1	2	3	4	5
4	2	3	4	5	6
6	3	4	5	6	7
8	4	5	6	7	8

La distribución de probabilidad y los cálculos de $E(\xi)$ y $E(\xi^2)$:

ξ	0	1	2	3	4	5	6	7	8	
$P(\xi)$	1/25	2/25	3/25	4/25	5/25	4/25	3/25	2/25	1/25	Σ
$\xi P(\xi)$	0	2/25	6/25	12/25	20/25	20/25	18/25	14/25	8/25	4
$\xi^2 P(\xi)$	0	2/25	12/25	36/25	80/25	100/25	108/25	98/25	64/25	20

- $E(\xi) = \mu_\xi = 4$ y $V(\xi) = 20 - 4^2 = 4 \rightarrow \sigma_\xi = 2$

- $\mu_\xi = \mu = 4$ y $\sigma_\xi = \sigma / \sqrt{n} = 2.8284 / \sqrt{2} = 2$

Problema 2: Igual que en el problema 1, $\mu = 4$ y $\sigma = 2.8284$. El número de elementos del espacio muestral, sin sustitución, es: ${}_5C_2 = 10$. En la siguiente tabla se muestran las medias de estas 10 duplas.

	0	2	4	6	8
0	--	1	2	3	4
2	--	--	3	4	5
4	--	--	--	5	6
6	--	--	--	--	7
8	--	--	--	--	--

La distribución de probabilidad y los cálculos de $E(\xi)$ y $E(\xi^2)$, son:

ξ	1	2	3	4	5	6	7	
$P(\xi)$	1/10	1/10	2/10	2/10	2/10	1/10	1/10	Σ
$\xi P(\xi)$	1/10	2/10	6/10	8/10	10/10	6/10	7/10	4
$\xi^2 P(\xi)$	1/10	4/10	18/10	32/10	50/10	36/10	49/10	19

- $E(\xi) = \mu_\xi = 4$ y $V(\xi) = 19 - 4^2 = 3 \rightarrow \sigma_\xi = 1.732$

$$\mu_\xi = \mu = 4 \quad \text{y} \quad \sigma_\xi = (\sigma / \sqrt{n}) \sqrt{\frac{N-n}{N-1}} = (2.8284/\sqrt{2}) \sqrt{\frac{5-2}{5-1}} = 1.732$$

Problema 3: Se tiene $\mu = 1.70$ y $\sigma = 0.15$

- Para n = 100: $\mu_\xi = \mu = 1.70$ y $\sigma_\xi = \sigma / \sqrt{n} = 0.15 / \sqrt{100} = 0.015$

  $P(X \geq 1.73) = 1 - P(Z < (1.73 - 1.70) / 0.015) = 1 - 0.9772 = 2.28\%$

- Para n = 50: $\mu_\xi = \mu = 1.70$ y $\sigma_\xi = \sigma / \sqrt{n} = 0.15 / \sqrt{50} = 0.021$

  $P(X \geq 1.73) = 1 - P(Z < (1.73 - 1.70) / 0.021) = 1 - 0.9236 = 7.64\%$

- Para n = 25: $\mu_\xi = \mu = 1.70$ y $\sigma_\xi = \sigma / \sqrt{n} = 0.15 / \sqrt{25} = 0.03$

  $P(X \geq 1.73) = 1 - P(Z < (1.73 - 1.70) / 0.03) = 1 - 0.8413 = 15.87\%$

- Para n = 1: $\mu_\xi = \mu = 1.70$ y $\sigma_\xi = \sigma / \sqrt{n} = 0.15 / \sqrt{1} = 0.15$

  $P(X \geq 1.73) = 1 - P(Z < (1.73 - 1.70) / 0.15) = 1 - 0.5793 = 42.07\%$

Problema 4: Se tiene $\mu = 350$ y $\sigma = 25$

Se determina el peso a partir del cual están los 400 novillos más pesados del hato, así: $P(X \geq x) = 400/2000 = 0.20 = P(Z \geq (x - 350) / 25)$, esto es equivalente a: $P(Z < (x - 350) / 25) = 0.8$, por lo tanto, $(x-350) / 25 = 0.7881$ y en consecuencia, $x = 369.7$

- Para n = 16: $\mu_\xi = \mu = 350$ y $\sigma_\xi = \sigma / \sqrt{n} = 25 / \sqrt{16} = 6.25$

  $P(X \geq 369.7) = 1 - P(Z < (369.7 - 350) / 6.25) = 1 - 0.9992 = 0.08\%$

- Para n = 4: $\mu_\xi = \mu = 350$ y $\sigma_\xi = \sigma / \sqrt{n} = 25 / \sqrt{4} = 12.5$

  $P(X \geq 369.7) = 1 - P(Z < (369.7 - 350) / 12.5) = 1 - 0.9418 = 5.82\%$

Problema 5: Se calculan $E(\xi) = \mu_\xi = 2.03$ y $V(\xi) = 4.615 - 2.03^2 = 0.4941 \rightarrow \sigma_\xi = 0.703$, luego, para la población: $\mu = 2.03$ y $\sigma = 0.703 \times \sqrt{2} = 0.994$

- Para la población (N), la $P(X \leq 0) = P(Z \leq (0 - 2.03) / 0.994) = 4/ N = 0.0207$, por lo tanto, $N \approx 193$.

- El total de muestras de dos trabajadores es: ${}_{193}C_2 = 18528$, de las cuales, el 32% (1% + 2% + 8% + 21%) tienen 1.5 o menos abuelos vivos en promedio, por lo tanto, $18528 \times 32\% \approx 5929$

Problema 6: Se tiene $\mu = 50$ y $\sigma = 1.75$

Para n = 3: $\mu_\xi = \mu = 50$ y $\sigma_\xi = \sigma / \sqrt{n} = 1.75 / \sqrt{3} = 1.01$

- $P(\xi \leq 50 + (3 \times 1.75)) - P(\xi \leq 50 - (3 \times 1.75)) = P(\xi \leq 55.25) - P(\xi \leq 44.75) = P(Z < 5.198) - P(Z < -5.198) = 100\%$

- $P(\xi \leq 50 + (2 \times 1.75)) - P(\xi \leq 50 - (2 \times 1.75)) = P(\xi \leq 53.5) - P(\xi \leq 46.5) = P(Z < 3.46) - P(Z < -3.46) = 100\%$

- $P(\xi \leq 50 + (1 \times 1.75)) - P(\xi \leq 50 - (1 \times 1.75)) = P(\xi \leq 51.75) - P(\xi \leq 48.25) = P(Z < 1.73) - P(Z < -1.73) = 91.64\%$

- $P(\xi \leq 50 + (0.5 \times 1.75)) - P(\xi \leq 50 - (0.5 \times 1.75)) = P(\xi \leq 50.875) - P(\xi \leq 49.125) = P(Z < 0.867) - P(Z < -0.867) = 61.56\%$

Problema 7: Se tiene $\mu = 50$ y $\sigma = 1.75$

Para n = 1: $\mu_\xi = \mu = 50$ y $\sigma_\xi = \sigma / \sqrt{n} = 1.75 / \sqrt{1} = 1.75$

$1 - P(49.5 \leq \xi \leq 51) = 1 - P(Z \leq 0.57) + P(Z \leq -0.29) = 89.84\%$

Problema 8: Se tiene $\mu_p = p = 0.025$ y $\sigma_p = \sqrt{\frac{pq}{n}} = 0.0221$

$P(X \geq 0.04) = P(Z \geq (0.04 - 0.025) / 0.0221) = 1 - P(Z \leq 0.68) = 24.83\%$

Problema 9: Se tiene $\mu_p = p = 0.025$ y $\sigma_p = \sqrt{\frac{pq}{n}} = 0.018$

$P(X \geq 0.04) = P(Z \geq (0.04 - 0.025) / 0.018) = 1 - P(Z \leq 0.83) = 20.33\%$

Problema 10: Se tiene $\mu_p = p = 0.02$ y $\sigma_p = \sqrt{\frac{pq}{n}} = 0.0044$

- $P(0.015 \leq X \leq 0.025) = P(Z \leq 1.14) - P(Z \leq -1.14) = 74.58\%$
- $P(X \leq 0.01) = P(Z \leq -2.27) = 1.16\%$
- $P(Z \leq (0.01 - 0.02) / \sqrt{\frac{0.0196}{n}}) = 10\% \rightarrow -0.01 / \sqrt{\frac{0.0196}{n}} = -1.28$, por lo tanto, $n \approx 321$

AUTO EVALUACIÓN 1

1. Considere la siguiente situación, se tiene una urna que contiene 28 fichas marcadas con las letras del alfabeto. Se extraen al azar una a una 10 fichas sin sustitución y se alinean en el orden de salida.

1.1. Justifique el por qué el anterior proceso constituye un experimento
1.2. Cuál es el espacio muestral de este experimento?
1.3. Cuánto elementos tiene el espacio muestral del experimento?
1.4. Cuál es la probabilidad de obtener la palabra "MURCIÉLAGO" en una secuencia de extracciones?

2. tres ruletas están marcadas con los colores rojo, verde, azul y blanco. Considere el experimento de girarlas simultáneamente. Si estamos interesados en la secuencia de colores en que caen las ruletas.

2.1. Determine el espacio muestral de este experimento
2.2. Si se define el suceso: Las tres ruletas caen en el mismo color. De cuántas maneras puede ocurrir el suceso?
2.3. Cuál es la probabilidad de que las ruletas caigan en colores diferentes?
2.4. Si se define como resultado de interés el número de rojos en que caen las ruletas. Determine el espacio muestral del experimento.
2.5. Justifique por qué el resultado de del punto anterior es una variable aleatoria. Qué valores puede tomar?
2.6. Determine la distribución de probabilidad de la variable aleatoria del punto anterior.
2.7. Cuál es la probabilidad de que dos de las ruletas caigan en el color rojo?

3. Un lote de 10 interruptores tiene 3 unidades defectuosas. Calcule las probabilidades de:

3.1. Obtener un defectuoso al extraer al azar uno de ellos del lote
3.2. No obtener defectuosos al extraer dos unidades del lote
3.3. Obtener tres defectuosos al extraer 10 unidades del lote
3.4. Justifique la distribución que usó para calcular las probabilidades pedidas en los tres puntos anteriores.

4. Un contenedor tiene un gran número de canicas de las cuales el 9% son verdes.

4.1. Cuál es la probabilidad de obtener exactamente 3 canicas verdes al extraer 10 de ellas al azar del contenedor?

4.2. Cuál es la probabilidad de obtener hasta 3 canicas verdes al extraer 10 de ellas al azar del contenedor?
4.3. Justifique la distribución que usó para calcular las probabilidades pedidas en los dos puntos anteriores.

5. El peso de los habitantes de una comunidad se distribuye normal con media de 48 kilogramos y desviación estándar de 5 kilogramos.

5.1. Cuál es la probabilidad de que un miembro de la comunidad elegido al azar pese menos de 35 kilogramos?
5.2. Qué porcentaje de la población pesa más de 60 kilogramos?
5.3. Qué porcentaje de la población pesa menos de 30 kilogramos?
5.4. Cuál es la probabilidad de que un miembro de la comunidad elegido al azar pese exactamente 45 kilogramos?

6. En relación con el problema 5, suponga que se registra el peso promedio de todas las posibles muestras de 50 miembros extraídas al azar.

6.1. Explique por qué estos pesos promedios constituyen una variable aleatoria.
6.2. Determine el tipo de distribución y los parámetros de esta variable aleatoria
6.3. Cuál es la probabilidad de que una de estas muestras tenga un peso promedio entre 38 y 55 kilogramos?

NOTAS:

- Interprete los resultados obtenidos en cada problema
- El tiempo sugerido de realización de esta prueba es de 4 horas.
- Evalúe su desempeño en la prueba y establezca claramente las dificultades encontradas.
- Revise el material en aquellos tópicos en los que encontró dificultades.

AUTO EVALUACIÓN 2

1. Los números de las matriculas del parque automotor de una ciudad se componen de cinco dígitos (del 0 al 9).

   1.1. Cuántos autos pueden matricularse en la ciudad?
   1.2. Determine el espacio muestral de este experimento.
   1.3. Si las matriculas se asignan al azar. Cuál es la probabilidad de que a un auto le corresponda un número de matricula cuyos 5 dígitos sean iguales?

2. Una máquina de soldadura aplica tres puntos de soldadura simultáneamente a láminas metálicas. Los puntos de soldadura se clasifican en buenos (b) y defectuosos (d). El operario de la máquina registra la calidad (b o d) de los puntos de soldadura en cada lámina.

   2.1. Discuta si los datos registrados por el operario en una jornada de trabajo constituyen un experimento. Especifique qué condiciones debe asumir.
   2.2. Determine el espacio muestral de este experimento
   2.3. Si se define el suceso: Dos de los tres puntos resultan buenos (b). De cuántas maneras puede ocurrir el suceso?
   2.4. Si se define como resultado de interés el número de puntos defectuosos por lámina. Cuál es el espacio muestral del experimento.
   2.5. Justifique por qué el resultado de del punto anterior es una variable aleatoria. Qué valores puede tomar?

3. Un lote de 10 interruptores tiene 3 unidades defectuosas. Calcule las probabilidades de:

   3.1. Obtener un defectuoso al extraer al azar uno de ellos del lote
   3.2. No obtener defectuosos al extraer dos unidades del lote
   3.3. Obtener tres defectuosos al extraer 10 unidades del lote
   3.4. Justifique la distribución que usó para calcular las probabilidades pedidas en los tres puntos anteriores.

4. Un contenedor tiene un gran número de canicas de las cuales el 9% son verdes.

   4.1. Cuál es la probabilidad de obtener exactamente 3 canicas verdes al extraer 10 de ellas al azar del contenedor?
   4.2. Cuál es la probabilidad de obtener hasta 3 canicas verdes al extraer 10 de ellas al azar del contenedor?
   4.3. Justifique la distribución que usó para calcular las probabilidades pedidas en los dos puntos anteriores.

5. El peso de los habitantes de una comunidad se distribuye normal con media de 48 kilogramos y desviación estándar de 5 kilogramos.

5.1. Cuál es la probabilidad de que un miembro de la comunidad elegido al azar pese menos de 35 kilogramos?
5.2. Qué porcentaje de la población pesa más de 60 kilogramos?
5.3. Qué porcentaje de la población pesa menos de 30 kilogramos?
5.4. Cuál es la probabilidad de que un miembro de la comunidad elegido al azar pese exactamente 45 kilogramos?

6. En relación con el problema 5, suponga que se registra el peso promedio de todas las posibles muestras de 50 miembros extraídas al azar.

6.1. Explique por qué estos pesos promedios constituyen una variable aleatoria.
6.2. Determine el tipo de distribución y los parámetros de esta variable aleatoria
6.3. Cuál es la probabilidad de que una de estas muestras tenga un peso promedio entre 38 y 55 kilogramos?

NOTAS:

- Interprete los resultados obtenidos en cada problema
- El tiempo sugerido de realización de esta prueba es de 4 horas.
- Evalúe su desempeño en la prueba y establezca claramente las dificultades encontradas.
- Revise el material en aquellos tópicos en los que encontró dificultades.

Fundamentos Estadísticos para el Control Estadístico de Calidad
Hernán Darío Rendón C.
Universidad Nacional de Colombia, Sede Medellín
Centro Editorial Facultad de Minas
Se utilizaron las fuentes Arial para el cuerpo de texto y para títulos.

125 AÑOS
1887 - 2012
Ingeniería con Trabajo y Rectitud

UNIVERSIDAD
NACIONAL
DE COLOMBIA
SEDE MEDELLÍN
FACULTAD DE MINAS