UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

COLEGIO DE CIENCIAS Y HUMANIDADES

SECRETARÍA ACADÉMICA

GUÍA DE ESTUDIO PARA:

EXAMEN DE CONOCIMIENTOS DISCIPLINARIOS

ESTADÍSTICA Y PROBABILIDAD I-II

Promoción XLI

ENERO 2020

Índice

PRESENTACIÓN

En esta guía se presenta una colección de problemas que ejemplifican los contenidos que debe dominar un profesor de la materia Estadística y Probabilidad en el CCH, así como una bibliografía que puede ser útil para preparar el examen de conocimientos y habilidades.

Los propósitos de la materia son:

- *El alumno interpretará formalmente resultados estadísticos, clarificando el papel del azar y valorando la variabilidad, con la finalidad de que verifique la importancia de la estadística y la probabilidad en la construcción de conocimientos y en la evaluación de hechos en diversos campos del saber, a partir del diseño y aplicación de un proceso de investigación estadística que incluya la formulación de preguntas, el levantamiento datos y el análisis de la información.*
- *El alumno desarrollará su habilidad para razonar usando un pensamiento estadístico, lo que le permitirá tomar decisiones sustentadas, juzgar críticamente la validez o pertinencia de la información estadística y realizar inferencias formales.*

Para que un profesor pueda conducir a sus alumnos al logro de estos objetivos, es necesario que conozca bien el programa de estudios de la materia, que aplique el enfoque que se le propone y que maneje bien no sólo los conocimientos incluidos en dicho programa, sino un contenido más amplio y profundo que le permita contestar dudas y motivar a sus estudiantes. En ese sentido, el examen de conocimientos que realizarán los aspirantes a impartir esta disciplina en el CCH contemplará:

I. Dominio de los conceptos y los métodos básicos de la Estadística y la Probabilidad.
II. Capacidad para resolver problemas aplicando adecuadamente dichos métodos y procedimientos.

El examen estará integrado por problemas similares a los que se proponen en esta guía y el aspirante deberá obtener una calificación mínima de 6 para acreditar esta parte de su evaluación, disponiendo de 3 horas para ese objetivo.

Es fundamental la presentación ordenada de los procedimientos que conduzcan a las respuestas o conclusiones solicitadas en el examen. El aspirante requiere contar con una **calculadora** y con la **tabla de la distribución normal estándar** que esté acostumbrado a utilizar.

ESTADÍSTICA Y PROBABILIDAD I

UNIDAD 1. OBTENCIÓN, DESCRIPCIÓN E INTERPRETACIÓN DE INFORMACIÓN ESTADÍSTICA.

Conocimientos básicos

a) Azar y probabilidad
b) Las variables estadísticas y su clasificación
c) Tablas de distribución de frecuencias de distintos tipos, incluyendo frecuencias absolutas, relativas y acumuladas
d) Distintos tipos de representaciones gráficas
e) Medidas de tendencia central y sus propiedades
f) Medidas de dispersión y sus propiedades
g) Medidas de posición y sus propiedades
h) Regla empírica para distribuciones aproximadamente simétricas

Procedimientos básicos

a) Construcción de tablas de distribución de frecuencias para representar el comportamiento de variables cualitativas y cuantitativas. Inferencias informales con base en esa información.
b) Construcción de representaciones gráficas: gráficas de barras, circulares y de caja, histograma y polígono de frecuencias, ojivas.
c) Interpretación de gráficas para describir el comportamiento de un conjunto de datos. Inferencias informales con base en la información que aportan.
d) Cálculo e interpretación de medidas de tendencia central y de dispersión en datos agrupados y no agrupados
e) Cálculo e interpretación de medidas de posición
f) Cálculo e interpretación del coeficiente de variación
g) Aplicación de la Regla Empírica

Problemas

1) Definir los siguientes conceptos:
   a) Población y muestras
   b) Variables estadísticas y datos
   c) Formas de recopilación de datos (levantamiento y experimentación)
   d) Ramas de la estadística

2) Describir las etapas de una investigación estadística.

3) Describir la clasificación de las variables estadísticas y detallar las características de cada tipo. Analizar ejemplos que pueden ser confusos como el código postal.

4) Los siguientes datos corresponden a la antigüedad docente (en años y décimas de año) de un grupo de profesores universitarios.

a) Construir una tabla de frecuencias adecuada a la información.

b) Construir la gráfica que se considere adecuada y describir las características principales de la información brindada por los datos.

c) Calcular las medidas de tendencia central e interpretar cada una de esas medidas en el contexto.

d) Escribir un pequeño reporte de entre cinco y diez renglones acerca del comportamiento de la variable.

14.7	11.6	35.8	12.2	8.7	11.6	34.8	38.8	23.2	5.1
29.0	21.2	12.9	13.0	17.8	34.9	43.3	22.9	15.3	26.3
8.6	11.8	6.2	15.8	4.1	5.9	11.8	12.2	14.5	47.5
7.0	48.1	14.5	11.0	34.9	15.8	31.5	14.5	45.3	18.5
9.0	4.3	5.8	16.7	25.8	28.0	24.9	15.5	18.9	27.3
11.2	34.3	12.4	22.5	20.4	30.4	30.4	33.4	22.5	36.4

5) a) Si el salario medio de los empleados de sexo masculino de la compañía A excede el de todos los empleados de sexo masculino de la compañía B y el salario medio de las empleadas de sexo femenino de la compañía A es mayor que el de todos los empleados de sexo femenino de la compañía B, ¿se deduce que el salario medio de todos los empleados de la compañía A excede el de todos los empleados de la compañía B? Explicar bien la respuesta.

b) ¿Puede ocurrir que la mayoría de los obreros de una empresa ganen menos del salario medio? Si es así, ¿en qué situación puede darse este caso? Explicar.

6) Las 5 primeras desviaciones de un conjunto de seis observaciones con respecto a la media, son – 2, 3, 7, 4 y –1.

a) ¿Cuál es el valor de la sexta desviación con respecto a la media?

b) Proporcionar una colección de datos que tenga las desviaciones anteriores con respecto a su media. ¿Cuántas muestras pueden construirse con esta característica?

c) Calcular la varianza y la desviación estándar de los datos proporcionados.

7) Construir colecciones de datos con las siguientes características.

a) 10 datos distintos que tengan media 20 y una desviación estándar menor a 1.5

b) 10 datos distintos que tengan media 20 y una desviación estándar entre 1.5 y 2.5

c) 10 datos distintos que tengan media 20 y una desviación estándar mayor a 2.5

8) El salario medio semanal de los empleados de una empresa es $5500 con desviación estándar de $300. La empresa otorgará un incremento salarial para lo cual analiza las siguientes alternativas:

A. Efectuar un aumento de $800 a cada empleado.
B. Incrementar en un 6% todos los salarios.
C. Un incremento salarial del 4% más un aumento fijo de $600 a cada empleado.
D. Aumentar en $500 todos los salarios más un 5% sobre el salario que resulte después de este aumento.

a) Calcular los salarios medios tras aplicar cada una de las alternativas anteriores.
b) ¿En cuál alternativa el salario medio aumenta más? ¿En cuál aumenta menos?
c) Calcular la varianza y la desviación estándar al aplicar cada una de las alternativas.
d) ¿En alguna opción disminuye la desviación estándar? ¿En cuál?

9) Los siguientes datos representan las edades en años de actores y actrices ganadores de un premio.

Actores:

32	37	36	32	51	53	33	61	35	45	55
47	39	76	37	42	40	32	60	38	56	48
40	43	62	43	42	44	41	56	39	46	48

Actrices:

50	44	35	80	26	28	41	21	61	38	49
33	74	30	33	41	31	35	41	42	37	26
34	35	26	61	60	34	24	30	37	31	34

Comparar el comportamiento de estos dos conjuntos de datos usando coeficientes de variación y diagramas de caja. ¿Qué conclusiones se pueden desprender? Explicar.

10) De los vuelos realizados por una compañía aérea A, el 20% no presentaron retraso. La distribución de frecuencias de los vuelos que sí presentaron retraso es la siguiente:

Retraso (min)	Número de vuelos
5	1000
10	800
15	650
20	350
30	200

a) ¿Cuál es el tiempo mediano de retraso de los vuelos con retraso? ¿Y del total de vuelos de la compañía?
Los siguientes incisos se refieren al total de vuelos.
b) Si el último vuelo realizado se encuentra entre el 20% de los vuelos con más retraso, ¿cuál es el mínimo retraso que puede haber presentado?

c) ¿Cuál es el tiempo de retraso que corresponde al percentil 32? ¿Y al percentil 64?

11) Se analizó el tamaño de las piezas de pan de centeno distribuidas en tiendas locales por cierta pastelería y al representar gráficamente la distribución del tamaño de estos panes se encontró un histograma más o menos simétrico con forma parecida a una campana. La longitud promedio del pan de 30 cm y la desviación estándar es de 1.3 cm. Aproximadamente ¿qué porcentaje de las piezas de pan tendrán un tamaño de:

a) Más de 31.3 cm?
b) A lo más de 27.4 cm?
c) Más de 27.4 cm y menos d 32.6 cm?

UNIDAD 2: OBTENCIÓN E INTERPRETACIÓN DE INFORMACIÓN ESTADÍSTICA CON DATOS BIVARIADOS

Conceptos básicos

a) Tablas de contingencia
b) Correlación lineal simple
c) Regresión lineal
d) Método de mínimos cuadrados

Procedimientos básicos

a) Construcción de tablas de contingencia para representar la relación entre dos variables cualitativas
b) Interpretación de la información contenida en las tablas de contingencia
c) Construcción de diagramas de dispersión para representar la relación entre dos variables cuantitativas
d) Cálculo e interpretación de los valores estimados de la pendiente y la ordenada al origen de la recta de mínimos cuadrados
e) Graficación de la recta de regresión
f) Cálculo e interpretación del coeficiente de correlación lineal simple
g) Utilización de la ecuación de la recta de mejor ajuste para predecir valores de la variable de respuesta

Problemas

12) Definir los siguientes conceptos:

a) Análisis de regresión.
b) Regresión lineal simple y múltiple
c) Coeficiente de correlación de Pearson

13) Se realizó una encuesta a 500 estudiantes de bachillerato acerca del tabaquismo y el género. Se obtuvo la siguiente información: 205 de los encuestados nunca han fumado, 124 de los cuales son mujeres; 65 dejaron de fumar y hay 42 mujeres entre ellos; los demás encuestados fuman y 60% de ellos son hombres.

a) Construir una tabla de contingencia. ¿Se aprecian elementos para pensar que hay alguna relación entre las variables de acuerdo a estos datos?
b) Calcular las frecuencias relativas (en porcentaje) y escribir al menos 4 observaciones sobre la información que brinda esta tabla.
c) Calcular los porcentajes por renglón y escribir al menos 4 observaciones sobre la información que brinda esta tabla.
d) Calcular los porcentajes por renglón y escribir al menos 4 observaciones sobre la información que brinda esta tabla.

14) Los datos siguientes son los tiempos de secado de un barniz y la cantidad de aditivo químico adicionado:

Cantidad de aditivo (gramos)	1	2	3	4	5	6	7	8
Tiempo de secado (horas)	7.2	6.6	4.5	3.7	4.6	4.2	4.0	3.3

a) Construir el diagrama de dispersión y hacer una conjetura sobre la posible relación lineal entre las variables.
b) Determinar la ecuación de la recta de regresión (por mínimos cuadrados).
c) Trazar la recta de regresión en el mismo plano cartesiano del diagrama de dispersión.
d) ¿Cuál es el tiempo de secado aproximado si se utilizan 3.5 gramos de aditivo? ¿Es confiable este resultado? Justificar la respuesta.
e) ¿Puede emplearse la recta de regresión para aproximar el tiempo de secado al usar 10 gramos de aditivo? Explicar la respuesta.
a) ¿Qué representa la ordenada al origen en el contexto de este problema? ¿Cómo puede interpretarse la pendiente de la recta en el contexto?

15) En un estudio sobre los hábitos de dar propina, se encontró que la ecuación de la recta de mejor ajuste es: $y = 0.5 + 0.177x$, donde x es el monto de la cuenta del restaurante y y es la cantidad de propina que deja el cliente. Se consideraron valores de x en el intervalo [10, 500].

a) ¿Qué representa la pendiente de esta recta? Explicar el sentido de su valor en el contexto.

b) ¿Qué representa la ordenada al origen? Explicar el sentido de su valor en el contexto.

c) Si la cuenta del restaurante fuera de $ 50.00, ¿qué pronosticaría la recta de mejor ajuste para la propina? ¿Y si fuera de $500?

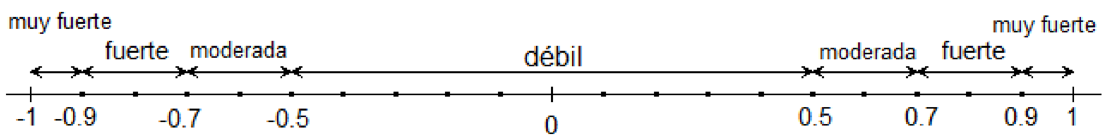
d) ¿Se puede usar la recta para pronosticar la propina en una cuenta de un monto de $1000? Explicar la respuesta.

16) Los siguientes datos representan una muestra del consumo de agua por día y la mayor temperatura ambiente para ese día.

Agua usada (millones de litros)	829	212	405	488	257	697	568	424
Temperatura máxima (grados Celsius)	39	4	25	26	10	36	32	24

b) Trazar el diagrama de dispersión y hacer una conjetura sobre qué tan lineal es la relación entre las variables, con base en los datos.

c) Explicar en qué consiste el análisis de correlación. Usar la siguiente escala para interpretar su valor:



d) Determinar qué valor tiene el coeficiente de correlación e interpretarlo en el contexto.

e) Encontrar el modelo lineal que mejor se ajusta a los datos.

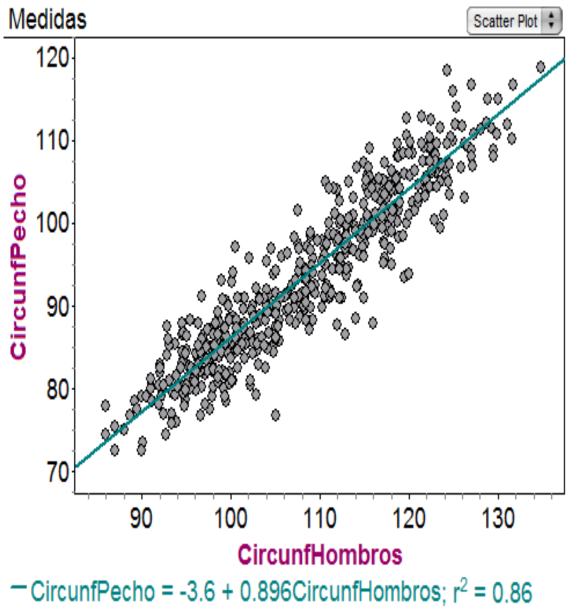
f) Traza la gráfica de la función anterior sobre el diagrama de dispersión.

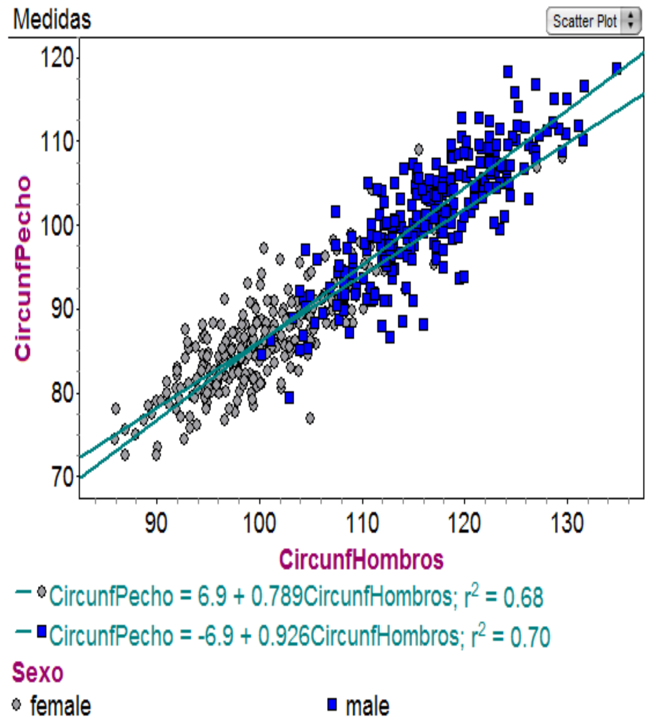
g) ¿Es confiable el modelo y el análisis obtenido? Justificar la respuesta.

17) a) Luego de que un estudiante calculó r para un conjunto de datos bivariados (apareados), descubrió, para su sorpresa, que la variable que se debería haber clasificado como x estaba clasificada como y, y la variable que se debería haber clasificado como y estaba clasificada como x. ¿Hay alguna razón para su sorpresa? ¿El análisis es equivalente para ambas elecciones de las variables $x - y$?

b) Haciendo una tarea, una estudiante obtuvo $S_{xx} = 145.22$, $S_{xy} = -210.58$ y $S_{yy} = 287.45$ para un conjunto de datos bivariados (apareados). Explicar por qué debe haber un error en sus cálculos.

18) En los siguientes diagramas de dispersión se muestra la relación entre la circunferencia de los hombros y la circunferencia del pecho para 507 personas adultas. En la segunda gráfica se muestran diferenciados los datos de mujeres y de hombres.





a) Escribe un reporte de entre cinco y diez renglones al respecto de las dos variables tomadas conjuntamente, interpretando la información brindada.
b) ¿A qué puede deberse la diferencia entre el comportamiento de las dos variables juntas para cada género?

UNIDAD 3. AZAR: MODELACIÓN Y TOMA DE DECISIONES

Conceptos básicos

a) Fenómenos aleatorios y deterministas
b) Regularidad estadística
c) Enfoques de la probabilidad: subjetivo, frecuencial, clásico
d) Eventos y espacio muestra
e) Eventos mutuamente excluyentes
f) Regla de la suma
g) Eventos independientes
h) Probabilidad condicional
i) Regla del producto

Procedimientos básicos

a) Construcción y descripción del espacio muestra de diversos eventos
b) Representación de eventos a partir de enunciados
c) Cálculo de probabilidades utilizando el enfoque frecuencial
d) Cálculo de probabilidades utilizando el enfoque clásico
e) Identificación y representación de eventos compuestos

f) Identificación y representación de eventos condicionados e independientes
g) Cálculo de probabilidades de los eventos descritos

Problemas

19) Se coloca una bola negra en la caja 1 y una bola roja en la caja 2. Se le pide a una persona que va pasando y que no sabe cómo fueron acomodadas las bolas, que saque la que se encuentra en la caja 2 y registre su color. ¿Qué clase de experimento es este: determinista o aleatorio? Explicar la respuesta.

20) Definir los siguientes conceptos:

a) Experimentos aleatorios y experimentos deterministas.
b) Espacio muestra de un experimento aleatorio y eventos.
c) Regularidad estadística o de frecuencias en los experimentos aleatorios.

21) En una urna hay 7 bolas, algunas son rojas y otras son negras. Se repite 50 veces el experimento de seleccionar una bola al azar y con reemplazo. En la siguiente tabla se reunió la información obtenida; R indica que fue roja y N indica que fue negra.

R	R	N	R	N	R	N	R	N	N
N	R	R	R	N	R	N	R	R	R
R	N	N	N	R	R	R	N	R	R
N	R	N	R	R	R	R	N	R	R
R	N	N	R	N	R	R	R	N	R

¿Cuántas bolas de cada color cree que tiene la urna? Escribir el procedimiento.

22) En una feria, se invita a los asistentes a participar en uno de dos juegos. Para saber por cuál decidirse, una persona observa, anota y ordena los resultados de dos muestras de 10 personas que han participado en cada juego, las perdidas (-) y premios (+) en efectivo que han obtenido las 20 personas se muestran en las siguientes listas.

Juego 1
15 -21 -4 50 -2 11 13 -25 16 -4
Juego 2
120 -120 60 -24 -21 133 -81 96 -132 18

a) Si se tiene la posibilidad de participar en un solo juego ¿cuál elegiría usted?
b) ¿En qué se ha basado la decisión?
c) ¿El número de observaciones permite una toma de decisiones bien fundamentada? Explicar.

23) El administrador de una tienda de regalos desea conocer la forma de pago de sus clientes, clasificados como clientes frecuentes y clientes eventuales. Para ello, revisó 600 pagos anteriores y obtuvo la siguiente información:

Cliente	Pago		Total
	Tarjeta	Contado	
Frecuente	180	165	345
Eventual	130	125	255
Total	310	290	600

Usar el enfoque frecuencial para aproximar la probabilidad de que el próximo cliente de la tienda

a) Sea eventual o pague con tarjeta.
b) Sea cliente frecuente y no pague de contado.
c) No pague con tarjeta ni sea eventual.
d) Sea cliente eventual y pague de contado.
e) Sea cliente frecuente o no pague con tarjeta.

24) Se lanzan dos dados comunes (cúbicos, marcados con los números del 1 al 6 y bien balanceados).

a) Escribir el espacio muestra.
b) Calcular la probabilidad de los siguientes eventos:
   A: sale al menos un 5.
   B: salen dos números que suman 8.
   C: salen dos números cuya diferencia absoluta es 3.
   D: el mayor de los dos números que salen, es 5.
c) Calcular la probabilidad de los eventos obtenidos aplicando las siguientes operaciones a los eventos del inciso b:
   $A \cup B$ $\quad$ $A \cap B$ $\quad$ $C \cap A^c$ $\quad$ $D \cap B^c$

25) Se forman números de 3 cifras eligiendo al azar y sin reemplazo, tres tarjetas de una colección de 6 tarjetas marcadas con los dígitos 1, 3, 4, 6, 7, 9. Calcular la probabilidad de los siguientes eventos:
E: El número formado es impar.
F: Se forma un número menor que 400
G: Se forma un número mayor que 499

26) Para formar un comité, se eligen al azar 4 personas de un grupo integrado por 5 mujeres y 6 hombres. Calcular la probabilidad de que el comité contenga:
A: únicamente mujeres.
B: únicamente personas del mismo sexo.
C: 2 mujeres y 3 hombres.

27) Cierta compañía tiene asignada la tarea de comparar la confiabilidad de dos estaciones de bombeo. Cada estación es susceptible a sufrir de dos tipos de fallas: falta de bombeo y fugas, cuando se presenta una de estas fallas o ambas, la estación debe quedar fuera de servicio. Los datos disponibles indican que se tienen las siguientes probabilidades:

Estación	Probabilidad de falla en bombeo	Probabilidad de fuga	Probabilidad de ambas fallas
1	0.09	0.12	0.06
2	0.07	0.10	0

¿Cuál estación tiene la mayor probabilidad de quedar fuera de servicio?

28) Se extraen dos bolas al azar y sin reemplazo de una urna que contiene 4 bolas rojas, 6 bolas negras y 8 blancas. Calcular la probabilidad de que:

a) La segunda bola extraída sea blanca dado que la primera fue roja.
b) La segunda bola extraída no sea negra dado que la primera sí fue negra
c) Las dos bolas extraídas sean blancas.
d) Las dos bolas extraídas sean del mismo color.

29) Sean A y B dos eventos de un mismo espacio muestra, tales que $P(A) = 0.35$, $P(B) = 0.45$ y $P(A \cup B) = 0.65$. Calcular

a) $P(A \cap B)$
b) $P(A \cap B^c)$
c) $P(A \cup B^c)$
d) $P(A \mid B)$
e) $P(B \mid A)$
f) $P(B^c \mid A)$
g) $P(A^c \cap B^c)$.

30) Se lanzan tres monedas comunes. Considerar los eventos:

A: por lo menos caen dos águilas
B: a lo más caen dos águilas
C: todos son soles o todos son águilas

De las parejas de eventos A y B, A y C, B y C, ¿cuáles son independientes y cuáles no lo son? Justificar la respuesta.

31) Si dos eventos son mutuamente excluyentes, ¿pueden también ser independientes? Explicar.

32) En cierta gasolinería, 45% de los clientes compran gasolina Magna, 30% compran diésel, y los restantes gasolina Premium. De los clientes que consumen gasolina Magna, sólo el 32% llena sus tanque~~s~~, el 55% de los que compran diésel llena su tanque, en tanto que el 45% de quienes llevan gasolina Premium llena su tanque. ¿Cuál es la probabilidad de que el cliente:

a) llene su tanque con gasolina Premium?
b) llene su tanque?
c) compre diésel si no llene su tanque?

ESTADÍSTICA Y PROBABILIDAD II

UNIDAD 1. MODELOS DE PROBABILIDAD Y SUS APLICACIONES

Conceptos básicos

a) Variable aleatoria
b) Variable aleatoria discreta y variable aleatoria continua
c) Distribuciones de probabilidad de variables aleatorias discretas
d) Parámetros: valor esperado y desviación estándar
e) Distribución Bernoulli
f) Distribución Binomial: experimento binomial, variable aleatoria binomial y parámetros
g) Distribución Normal: la distribución Normal como modelo continuo del comportamiento de una gran diversidad de fenómenos aleatorios; propiedades geométricas y analíticas; significado y ventajas de la estandarización.

Procedimientos básicos

a) Cuantificación de la probabilidad de eventos utilizando una variable aleatoria
b) Construcción de la distribución de probabilidad y la distribución de probabilidad acumulada asociadas a una variable aleatoria discreta
c) Representación gráfica de la distribución de probabilidad
d) Cálculo e interpretación del valor esperado y de la desviación estándar de una variable aleatoria
e) Cálculo de probabilidades en experimentos binomiales, del valor esperado y de la desviación estándar,
f) Cálculo de probabilidades utilizando la distribución normal estándar en problemas contextualizados, así como determinación del valor de la media o la desviación estándar dados los datos necesarios.

Problemas

33) En una caja hay 6 fusibles, dos de los cuales están defectuosos. Se van seleccionando fusibles al azar, de uno en uno y sin reemplazo, hasta que salen los dos defectuosos. La variable aleatoria X indica el número de extracciones realizadas.

a) ¿Qué valores puede tomar la variable X?
b) ¿Se trata de una variable aleatoria discreta o continua?
c) Describir la distribución de probabilidad (o función de densidad) de X.
d) Determinar la distribución acumulada (o función de distribución) de X.

34) Se lanza un dardo al azar sobre un blanco circular de 30 cm de radio. La variable aleatoria X indica la distancia del centro al punto de contacto.

a) ¿Qué valores puede tomar la variable X?
b) ¿Se trata de una variable aleatoria discreta o continua?
c) Determinar la distribución acumulada (o función de distribución) de X.

35) Las probabilidades de que un inspector de construcción observe 0, 1, 2, 3, 4 o 5 violaciones al reglamento de construcción en un edificio construido en la ciudad son, respectivamente 0.41, 0.22. 0.17, 0.13, 0.05 y 0.02.

a) Se elige un edificio al azar en la ciudad. Calcular la probabilidad de que X sea menor que 3 y la probabilidad de que se encuentre al menos una violación.
b) Determinar el número esperado de violaciones a dicho reglamento e interpretar su valor en el contexto.
c) Calcular la desviación estándar del número de violaciones e interpretar su valor.

36) En la siguiente tabla se muestran los valores de la variable aleatoria *X* que indica el número de automóviles vendidos por día en una agencia, de acuerdo a registros de 1000 días anteriores.

X	0	1	2	3	4	5	6	7
Núm. de días	49	87	148	234	297	135	30	20

a) Aproximar frecuencialmente la distribución de probabilidad de la variable aleatoria *X*.
b) Calcular la probabilidad de que mañana la agencia venda al menos 1 un auto pero menos de 5 autos.
c) Calcular la media y la desviación estándar de *X* e interpretar estos valores en el contexto.

37) Se lanza una moneda legal hasta que aparezca un sol o cinco águilas. Determinar el número esperado de lanzamientos de la moneda.

38) El teorema del binomio establece que si *a* y *b* son números reales y *n* es un número natural, entonces $(a+b)^n = \sum_{x=0}^{n} \binom{n}{x} a^x b^{n-x}$. Usa este resultado para verificar que la función

$$f(x) = \begin{cases} \binom{n}{x} a^x b^{n-x} & si\ \ x = 0, 1, 2, \dots n \\ 0 & en\ otro\ caso \end{cases}$$

es una función de densidad (o distribución de probabilidad).

39) Cada día, un tipo de acciones sube $1 con probabilidad 0.45 o baja $1 con probabilidad 0.55, independientemente del precio que haya tenido en cualquier otro día.

a) Calcular la probabilidad de que después de 6 días, una de estas acciones vuelva a su precio original.

b) Determinar el número esperado de aumentos en el precio de una de estas acciones durante los próximos 12 días, y la desviación estándar de ese número.

40) a) Escribir las condiciones que deben cumplirse para que se presente una distribución Binomial

b) Si en el problema anterior, la probabilidad de que la acción suba $1 o baje $1 cambia cada día, ¿se trata de una distribución binomial?

c) Si estas probabilidades dependen de lo que haya ocurrido los dos días previos, ¿se trata de una distribución binomial?

41) La probabilidad de obtener un falso positivo en cierta prueba de laboratorio, es 0.005 independientemente de lo que haya ocurrido en pruebas anteriores del mismo tipo.

a) Calcular la probabilidad de que en las próximas 10 pruebas de ese tipo
   - No ocurran falsos positivos.
   - Ocurra exactamente 1 falso positivo.
   - Ocurran más de un falso positivo.

b) Determina el número esperado de falsos positivos en las próximas 10 pruebas de ese tipo y su desviación estándar.

42) a) ¿Cómo explicarías a un grupo de alumnos de bachillerato, de manera sencilla y clara, que al hacer tender número n de pruebas a ∞, una variable con distribución binomial tiende a distribuirse como Normal?

b) Escribe un ejemplo en donde se use la aproximación Normal a una distribución Binomial.

c) Además de usarse como aproximación a una Binomial con n grande, ¿en qué tipo de situaciones se presentan distribuciones con forma de campana de Gauss?

43) Determina el valor o valores de la variable aleatoria Z que acoten:

a) El 65% del área al centro de la curva de la distribución normal (de manera simétrica respecto al eje Y).

b) El 25% en la cola izquierda de la curva de la distribución normal.
c) El 10% en la cola derecha de la curva de la distribución normal.
d) El 10% distribuido en ambas colas de la curva de manera simétrica

44) La producción diaria de una línea de ensamblaje se distribuye normalmente con una media de 165 unidades y una desviación estándar de 5.

a) Determinar la probabilidad de que el número de unidades producidas por día:
- Sea menor que 162 unidades.
- Sea mayor que 173 unidades.
- Esté entre 152 y 174 unidades.
- Esté entre 171 y 177 unidades.

b) Determinar el número de unidades x tal que en el 75 % de los días el número de unidades producidas rebase a x.

45) Las monedas de 50 centavos tienen pesos distribuidos normalmente con una media de 5.67 gramos y una desviación estándar de 0.07 gramos.

a) Si una máquina expendedora se ajusta de modo que rechace las monedas de 50 centavos que pesen menos de 5.53 gramos o más de 5.81 gramos, ¿qué porcentaje de monedas será rechazado?

b) Determinar cuánto pesan las monedas de 50 centavos aceptadas, si la máquina se reajusta de modo que rechace el 1.5% más ligero y el 1.5% más pesado.

46) Se sabe que los tiempos *en espera*, cuando se llama al *call center* de un banco determinado, están normalmente distribuidos con una desviación estándar de 1.3 minutos. Determinar el tiempo medio de espera si no más del 10% de quienes llaman esperan más de 6 minutos, como afirma el banco.

47) En un estudio de un gran número de personas adultas sanas, se encontró que 30% duermen menos de 7.2 horas diarias, mientras que 40% duermen menos de 7.5 horas diarias. Suponiendo que las horas de sueño se distribuyen normalmente, ¿cuáles son los valores de la media y la desviación estándar?

48) El tiempo de servicio de cierta marca de llantas de automóviles sigue una distribución normal con una media y una desviación estándar de 32 mil y mil kilómetros, respectivamente. Determina el porcentaje de llantas vendidas que se requiere reemplazar, si esta marca de llantas se garantiza por 30 mil kilómetros.

UNIDAD 2. ESTIMADORES E INTRODUCCIÓN A LA INFERENCIA ESTADÍSTICA

Conceptos básicos

a) Conceptos de Población Objetivo y Muestreo Aleatorio Simple
b) Parámetros, estadísticos y estimadores.

c) Concepto de variabilidad muestral
d) Los estadísticos como variables aleatorias (en particular la media muestral y la proporción muestral)
e) Distribución muestral de la media.
f) Distribución muestral de proporciones.
g) Teorema del Límite Central

Procedimientos básicos

a) Identificación de los conceptos de población objetivo y muestreo aleatorio simple
b) Selección de muestras de poblaciones finitas o infinitas.
c) Identificación de parámetros y estadísticos.
d) Graficación de las distribuciones muestrales para la media y la proporción.
e) Cálculo de los parámetros de las distribuciones muestrales y su relación con los parámetros de la población.
f) Comprensión e interpretación del Teorema de Límite Central.
g) Cálculo de probabilidades en las distribuciones de medias y proporciones muestrales dentro del contexto de un problema.

Problemas

49) a) Definir los conceptos de muestra aleatoria y muestra aleatoria simple.
b) ¿Cómo explicaría a estudiantes de bachillerato el doble sentido de las muestras aleatorias: como colección de variables aleatorias y como colección de valores (subconjunto de una población)?
c) Si una población tiene distribución normal con media μ y varianza σ^2, ¿cuál es la distribución de la media muestral $\overline{X}$ en una muestra de tamaño n?
d) Dar una explicación sencilla, breve y precisa del Teorema Central del Límite.

50) Un experimento consiste en lanzar una moneda legal 50 veces. Sea X la variable que indica el número de soles obtenidos.
a) ¿Cuál es la distribución real de X? ¿Cuál es su media y su desviación estándar?
b) ¿Por qué se puede usar una distribución normal para calcular probabilidades sobre X?
a) Si $\hat{p} = \frac{X}{50}$ representa la proporción de soles obtenidos, determinar $\mu_{\hat{p}}$ y $\sigma_{\hat{p}}$
b) Calcular $P(\hat{p} < 0.4)$ y $P(0.45 < \hat{p} < 0.63)$.

51) Suponga que la proporción de mujeres en el Colegio actualmente es de 0.55. Calcule el error estándar de la proporción de mujeres en una muestra de n alumnos seleccionados aleatoriamente, para n = 100, n = 200, n = 500 y n =

1000. ¿Qué se puede decir acerca del error estándar de la proporción al aumentar el tamaño de la muestra?

52) La mayoría de las personas creen que el desayuno es el alimento más importante del día pero 25% de los adultos no desayunan. Se elige una muestra aleatoria de 200 adultos.

a) Encontrar los parámetros y dibujar la distribución aproximada del número X de adultos de la muestra que desayunan.

b) ¿Cuál es la probabilidad de que el número de adultos de la muestra que desayunan, esté entre 140 y 150?

53) En una empresa se sabe que el 30% de los pedidos provienen de clientes nuevos. Si se selecciona una muestra aleatoria de 100 pedidos y se calcula la proporción de clientes nuevos en la muestra:

a) Dibuja la distribución de la proporción muestral.

b) ¿Cuál es la probabilidad de que la proporción muestral esté entre 0.20 y 0.40?

c) ¿Cuál es la probabilidad de que la proporción muestral sea mayor que 0.34?

54) La variable aleatoria X, que representa el número de cerezas en una empanada de cierta marca, tiene la siguiente distribución de probabilidad:

X	4	5	6	7
$P(x)$	0.2	0.4	0.3	0.1

a) Determinar la media y la varianza de X.

b) Determinar la media y la varianza de la media muestral para muestras aleatorias de 36 empanadas de cereza.

c) Calcular la probabilidad de que el número promedio de cerezas en una muestra de 36 empanadas sea menor que 5.5.

55) En una panadería se hornean diariamente piezas de pan con un peso medio de 482 gramos. Si se supone que la desviación estándar es de 18 gramos y se selecciona al azar una muestra de 40 piezas.

a) Calcular los parámetros de la distribución de la media muestral y dibujar la distribución aproximada (según el TCL).

b) ¿Cuál es la probabilidad de que la media muestral se encuentre entre 475 y 495 gramos?

c) ¿Cuál es la probabilidad de que la media muestral tenga un valor menor a 478 gramos?

d) ¿Cuál es la probabilidad de que la media muestral se encuentre a ± 5 gramos de la media poblacional?

56) Si una muestra aleatoria de tamaño 36 se extrae de una población con una media de 278 y si 86% de las veces que se selecciona una muestra aleatoria, la media muestral es menor a 280, ¿cuál es la desviación estándar de la población?

57) En los tabulados del INEGI se dan datos de la población total del país por Estado, edad y grupos quinquenales de edad y su distribución según sexo. Utiliza la información pertinente del sitio

http://www.conapo.gob.mx/es/CONAPO/Tabulados_basicos

para contestar lo siguiente si se eligiera al azar una muestra de 100 habitantes del país,

a) ¿Cuál es la probabilidad de que la media de edades de la muestra fuera mayor a 23 años?
b) ¿Cómo se distribuye la media de edades de la muestra?
c) Haga un dibujo de la distribución y sombree el área que ilustre el resultado que se obtuvo.

UNIDAD 3. INFERENCIA ESTADÍSTICA

Conceptos básicos

a) Estimación puntual y por intervalos para la media y la proporción poblacionales.
b) Nivel de confianza
c) Error de estimación
d) Tamaño de muestra
e) Prueba de hipótesis para la media y la proporción
f) Hipótesis estadísticas
g) Nivel de significancia

Procedimientos básicos

a) Elección de estimadores puntuales para la media y para la proporción poblacionales.
b) Construcción e interpretación de intervalos de confianza para la media y para la proporción de la población.
c) Cálculo del valor del tamaño de la muestra (***n***) para diferentes cotas de error y niveles de confianza,
d) Elaboración de hipótesis estadísticas para la media y para la proporción de la población.
e) Identificación de los tipos de errores asociados con la prueba de hipótesis,

f) Reconocimiento de los elementos que componen una prueba de hipótesis,
g) Realización de pruebas de hipótesis para la media y la proporción de la población en diferentes problemas, así como la interpretación de los resultados.

Problemas

58) a) Se tienen dos estadísticos que sirven como estimadores del mismo parámetro, uno insesgado y el otro sesgado. Explicar porque usualmente se preferirá el estimador insesgado.

b) ¿El que un estimador sea insesgado permite asegurar que es un buen estimador? ¿Por qué?

c) ¿Qué otras consideraciones es necesario tomar en cuenta?

d) Si se elige una muestra aleatoria de *n* elementos, ¿cuáles es el estimador usual de la media poblacional? ¿Es insesgado?

59) a) Definir lo siguiente al realizar una estimación por intervalos.

- Intervalos aleatorios e intervalos de confianza.
- Nivel de confianza en una estimación por intervalos.

b) Un investigador de mercado afirma que tiene un nivel de confianza del 95% en que la media de las ventas mensuales de un producto está entre $170,000 y $200,000. Explicar el significado de esta afirmación.

60) En una ciudad, el Departamento de Aguas toma una muestra aleatoria de 1,200 hogares y observa los medidores de agua. Se obtiene un consumo medio diario de agua en la muestra de 116 litros y una desviación estándar de 19.3 litros. Determinar un intervalo de confianza al 90% del consumo medio diario de agua en la ciudad e interprétalo en el contexto descrito.

61) Escribir las fórmulas para los extremos de un intervalo al estimar la media poblacional con un nivel de confianza de 1 - $\propto$, usando el estimador puntual $\overline{X}$, cuando se conoce la varianza poblacional y cuando no se conoce.

a) ¿Qué relación hay entre el nivel de confianza y la longitud del intervalo? Justificar la respuesta.

b) ¿Qué relación hay entre el tamaño de la muestra y la longitud del intervalo? Justificar la respuesta.

62) Antes de adquirir un embarque grande de carne molida, un fabricante de salchichas desea estimar el peso medio de la grasa por kilogramo con un nivel de confianza de 96%. Por experiencia, se sabe que la desviación estándar de la cantidad de grasa por kilo, es de 80 gramos. El comprador toma una muestra de *n* kilos de carne seleccionados al azar en distintas partes del embarque. ¿De qué tamaño debe ser *n* si desea que su error de estimación no sea mayor a 25 g?

63) Una muestra aleatoria de tamaño 300 de una población binomial produjo una media de 110 éxitos. Determina un intervalo de confianza de 95% para p e interprétalo.

64) En cierto día, se negociaron 2500 paquetes distintos de acciones. Una muestra aleatoria de 100 paquetes de acciones dejó ver que 70% bajaron de precio.

a) Estimar el error estándar de la proporción de acciones de la muestra que bajaron de precio.

b) Construir un intervalo de confianza del 99% para la proporción de acciones que fueron a la baja.

65) Se quiere estimar la proporción de conductores que exceden el límite de velocidad en una carretera ¿Qué tan grande debe ser la muestra para que el error de estimación sea a lo sumo 0.04, con niveles de confianza de

a) 90 %?

b) 94 %?

c) 98%

66) En una muestra aleatoria de 400 personas entrevistadas en una ciudad grande, 288 dijeron que se oponían a la construcción de más vías rápidas.

a) Construye un intervalo de confianza del 94% para la proporción correspondiente de la población en la que se efectúa el muestreo.

b) ¿Qué podemos decir con una confianza del 94 % acerca del error máximo, si usamos la proporción de la muestra como una estimación de la proporción de la población?

67) Definir los conceptos

a) Hipótesis estadística.

b) Pruebas de hipótesis basadas en una muestra aleatoria.

c) Los tipos de errores que se pueden cometer al realizar una prueba de hipótesis.

d) Tipos de juegos de hipótesis para un parámetro poblacional θ.

e) Nivel de significancia de una prueba de hipótesis.

f) Potencia de una prueba de hipótesis.

68) a) ¿Por qué es posible rechazar una hipótesis nula cuando es verdadera?

b) ¿Por qué es posible no rechazar una hipótesis nula cuando es falsa?

c) ¿Qué relación tiene α con el error tipo I?

d) ¿Qué relación tiene β con el error tipo II?

e) Para un tamaño de muestra dado, si α se reduce de 0.05 a 0.01, ¿qué pasa con β?

69) Encontrar las regiones de rechazo apropiadas para pruebas con muestras grandes en las que se usa el estadístico Z en los siguientes casos:

a) Una prueba de cola derecha con $\alpha = 0.01$.

b) Una prueba de dos colas con nivel de significancia de 5%

c) Una prueba de cola izquierda al nivel de significancia de 5%

70) Encontrar el valor-p para las siguientes pruebas con el estadístico Z usando muestras grandes.

a) Una prueba de cola derecha con z observada de 1.15

b) Una prueba de dos colas con z observada de -2.78

c) Una prueba de cola izquierda con z observada de -1.81

71) La cantidad de agua que se recomienda consumir al día, es de 8 porciones de 232 *ml*. Sin embargo, se han hecho estudios que reflejan que tomamos en promedio 4.6 pociones de 232 *ml*. Se seleccionó al azar una muestra de 100 estudiantes del Colegio y se observó su consumo de agua durante 24 horas. La cantidad media consumida por esta muestra fue de 1140 *ml* y su desviación estándar fue 325 *ml*. Con un nivel de significancia $\alpha = 0.05$, ¿hay suficiente evidencia de que los estudiantes del Colegio consumen en promedio más agua al día que la cantidad que consumen otras personas?

72) Una muestra aleatoria de 64 bolsas de palomitas de maíz con queso, pesan en promedio 148 gramos con una desviación estándar de 6.8 *g*. Probar la hipótesis de que $\mu = 156$ *g* contra la hipótesis alternativa $\mu < 156$ *g* con nivel de significancia de 0.04. Escribir una interpretación clara de la conclusión en el contexto.

73) El departamento de seguridad de una fábrica requiere saber si el promedio de tiempo real requerido por el guardián nocturno para hacer su ronda es 25 minutos o no. En una muestra aleatoria de 36 rondas el guardián nocturno promedió 25.8 minutos con una desviación estándar de 1.5 minutos. Determinar si ésta evidencia es suficiente para rechazar la hipótesis nula $\mu = 25$ minutos con un nivel de significancia de 0.01

74) Un fabricante afirma que por lo menos 20% del público prefiere su producto. Se toma una muestra de 100 personas para verificar su afirmación. Con $\boldsymbol{\alpha} = 0.05$, ¿qué tan pequeño debe ser el porcentaje de preferencia por ese producto en la muestra para poder refutar la afirmación del fabricante?

75) El fabricante de un limpiador de manchas afirma que su producto remueve 90% de todas las manchas. Si en una muestra aleatoria, el limpiador de manchas remueve 11 de 16 manchas, pruebe la hipótesis nula $p = 0.90$ contra la hipótesis alternativa $p < 0.90$ en el nivel de 0.05 de significancia.

76) En una universidad se estima que a lo más 25 % de los estudiantes se trasladan en bicicleta a la escuela. ¿Puede esta ser una estimación válida si en una muestra aleatoria de 90 estudiantes universitarios, se encuentra que 28 van en bicicleta a la escuela? Utilice un nivel de significancia de 0.05.

BIBLIOGRAFÍA

BÁSICA

1. Freund, J. (1994). *Estadística Elemental*, Octava edición, Prentice Hall, México.
2. Batanero, Carmen; Rodino, Juan D.(2001) *Análisis de datos y didáctica de la Estadística*. Grupo de Investigación en Educación en Estadística. Departamento de Didáctica de la Matemática. Universidad de Granada.
3. Castillo Padilla Juana, Gómez Arias Jorge. (1998). Estadística Inferencial Básica. México. Grupo Editorial Iberoamérica.
4. Chao, Lincoln L (1985). *Introducción a la Estadística.* CECSA, México.
5. Daniel, Wayne D. (1984). *Bioestadística:* 5ª reimpresión, Limusa, México.
6. Flores García, Rosalinda, Lozano de los Santos, Héctor, (1998) Estadística *Aplicada para Administración.* Grupo Editorial Iberoamérica, México.
7. Johnson, Robert. (1990). *Estadística Elemental,* Iberoamérica, México.
8. Mendenhal/Reinmuth. (1978). *Estadística para Administración y Economía,* Wadsworth International/ Iberoamérica, Massachusetts USA.
9. Stevenson, William J. (1985) *Estadística para Administración y Economía,* Trillas, México.
10. Triola, Mario F. (2000), *Estadística elemental*. Pearson Educación. México.

COMPLEMENTARIA

1. Battacharya, Gouri K y Johnson Richard A. (1977) *Statistical Concepts and Methods,* John Wiley & Sons N.Y., Santa Bárbara-London-Sydney-Toronto.
2. Canavos, George C. (1987). *Probabilidad y Estadística: Aplicaciones y Métodos,* McGraw-Hill, México.
3. Chou, Ya-Lun. (1975). *Análisis Estadístico, Segunda Edición,* Interamericana, México.
5. Díaz Godino, Juan, Batanero Bernabeu, María del C. y Cañizares Castellanos, María de Jesús. (1987) *Azar y Probabilidad: Fundamentos Didácticos y Propuestas Curriculares,* Madrid, Síntesis.
6. Dixon y Massey, (1965). *Introducción al Análisis Estadístico,* Mc-Graw-Hill, México.
7. Hacking, Ian. (19991). *La Domesticación del Azar.* Gedisa, Barcelona.
8. Lohr, Sharon L. (2000) *Muestreo; Diseño y Análisis*. Thomson Editores, México.
9. Mendenhall, William, Sheaffer, Richard l. y wackerly Dennis o. (1986). *Estadística Matemática con Aplicaciones.* Iberoamérica. México.
10. Said Infante, Gil y Zárate de Lara, Guillermo P. (1984). *Métodos Estadísticos: Un enfoque Interdisciplinario,* Trillas. México.
11. Sheaffer, Mendenhall y ott. (1986). *Elementos de Muestreo.* (1986) Iberoamérica. México.
12. Vilenkin, N. (1972) *¿De Cuántas Formas?: Combinatoria,* Mir Moscú, URSS, 1972.
13. Yamane, Taro. *Estadística.* Tercera edición, Harla, México.