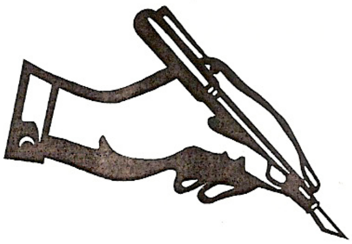
UNIVERSIDAD AUTONOMA DE MADRID

EEES: Plataforma Moodle Estadística Teórica II
Proyecto CEyE-L2/2
Facultad de Ciencias Económicas y Empresariales
Departamento de Economía Aplicada
Profesor: Santiago de la Fuente Fernández

Estadística Teórica II



e-learning UAM

1. En el cálculo de probabilidades de una variable discreta y continua ¿existen diferencias). Razone la respuesta.

Respuesta: Para una variable discreta $X = \{x_i\}_{i=1,2,\cdots,n}$ existe una función de probabilidad $P(X = x_i)$, verificándose que $\sum_{i=1}^{n} P(X = x_i) = 1$. La función de distribución

$F(x) = P(X \leq x) = \sum_{x_i \leq x} P(X = x_i) = 1$.

Para una variable continua X existe una función de densidad f(x) tal que la función de distribución

$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\,dx$

2. Defina el concepto de variable aleatoria.

Respuesta: Sea P la función de probabilidad definida para los sucesos de un determinado experimento aleatorio, siendo Ω su espacio muestral.

Sea una función que a cada suceso elemental le hace corresponder un número real

$$\begin{aligned} X : \Omega &\longrightarrow R \\ w &\longrightarrow X(w) \in R \end{aligned}$$

y representamos por $(X \leq a)$ el suceso $\{w \in \Omega \,/\, X(w) \leq a\}$.

Diremos que X es una variable aleatoria si existe la probabilidad $P(X \leq x)$ para todo número real x.

La función $F(x) = P(X \leq x)$ se denomina función de distribución.

3. ¿Qué relación existe entre la distribución Binomial y la de Bernouilli?

Respuesta:

La variable aleatoria X sigue una distribución de Bernouilli $B(1, p)$ si su función de probabilidad es:

$$P(X = x) = \begin{cases} p & \text{si } x = 1 \\ q = 1 - p & \text{si } x = 0 \\ 0 & \text{resto de valores de } x \end{cases} \qquad 0 < p < 1$$

La variable aleatoria X sigue una distribución binomial $B(n, p)$ si es suma de n variables de Bernouilli $B(1, p)$ independientes, su función de probabilidad será:

$$P(X = x) = \begin{cases} \binom{n}{x} p^x q^{n-x} & \text{si } x = 0, 1, 2, \cdots, n \\ 0 & \text{resto de valores de } x \end{cases}$$

4. Explique cuáles son las distribuciones asociadas a la normal y para que se utilizan.

Respuesta:

χ^2_n: Dadas n variables aleatorias independientes $X_i \approx N(0, 1)$, entonces la distribución de la variable $\sum_{i=1}^{n} X_i^2$ se denomina chi-cuadrado con n grados de libertad.

t_n : Si X es una variable aleatoria normal $N(0,1)$, e Y es una χ_n^2, ambas independientes, entonces la distribución de la variable $\frac{X}{\sqrt{\frac{Y}{n}}}$ se denomina t de Student con n grados de libertad.

$F_{n,m}$: Si X e Y son variables aleatorias independientes, siendo $X \approx \chi_n^2 \ N(0,1)$ e $Y \approx \chi_m^2$, entonces la distribución de la variable $\frac{X/n}{Y/m}$ se denomina F de Snedecor con n y m grados de libertad.

Tienen utilidad en los procesos de estimación pues son los modelos de distribución que siguen ciertos estimadores.

5. Explique el objetivo de la inferencia estadística.

Respuesta: La Inferencia Estadística persigue la obtención de conclusiones sobre un gran número de datos, basándose en la observación de una muestra obtenida de ellos; también intenta medir su significación, es decir la confianza que nos merecen.
Tiene como objetivo obtener información sobre el valor de algún parámetro poblacional, como por ejemplo la media o la varianza, o sobra la forma de la variable aleatoria, como por la ejemplo la función de densidad o de cuantía. Esta determinación se hace a partir de la información contenida en una muestra aleatoria.

6. Explique que es una muestra aleatoria simple.

Respuesta: Sea X una variable aleatoria de una población con función de distribución F(x).
Si las variables aleatorias $X_1, X_2, \cdots, X_n$ son independientes y tienen la misma función de distribución F(x) que la variable X, entonces se dice que las variables $X_1, X_2, \cdots, X_n$ constituyen una muestra aleatoria simple de tamaño n.

7. ¿Cuál es el objetivo de la estimación puntual?. Razone la respuesta.

Respuesta: Obtener el valor de un parámetro poblacional desconocido mediante la elección de un estadístico muestral (estimador). Obtenida la muestra, el valor del estimador se utilizará como valor del parámetro poblacional.

8. ¿Cuál es el objetivo de un contraste de hipótesis?. Razone la respuesta.

Respuesta: Es confirmar o rechazar una determinada hipótesis efectuada sobre el valor de un parámetro poblacional o sobre la distribución de una variable aleatoria. Para un determinado estadístico, determinaremos el conjunto de valores que admitimos para aceptar la hipótesis (región de aceptación) y el conjunto complementario (región de rechazo). Para efectuar el contraste, se elige una muestra y se calcula el valor que proporciona el estadístico en cuestión, lo que permite aceptar la hipótesis o bien rechazarla y, en su caso, aceptar una hipótesis alternativa.

9. Concepto de población, muestra y estadístico muestral. Ponga un ejemplo.

Respuesta:
POBLACIÓN.- Es el conjunto de individuos que son objeto del estudio estadístico, de forma que existe una variable aleatoria X asignando un número real a cada uno de los individuos de la población.
MUESTRA.- De tamaño n, es un conjunto de n observaciones de la variable aleatoria $(X_1, X_2, \cdots, X_n)$.
ESTADÍSTICO MUESTRAL.- Es una función de las variables $X_1, X_2, \cdots, X_n$ que componen la muestra.

EJEMPLO.- Una muestra puede ser el conjunto de clientes de unos grandes almacenes durante al año 2011, siendo X el gasto efectuado por cada cliente. La muestra queda formada por la elección de k clientes $(X_1, X_2, \cdots, X_k)$ que anotamos sus gastos. Un estadístico muestral, sería por ejemplo, la media muestral $\overline{x} = \frac{1}{n}\sum_{i=1}^{k} x_i$

10. Explique conceptualmente qué es el sesgo de un estimador y cuál es su interpretación.

Respuesta: Se denomina sesgo de un estimador $\hat{\theta}$ a la diferencia $E(\hat{\theta}) - \theta$, donde θ es el parámetro a estimar. Puede ser positivo, negativo o nulo. Una propiedad deseable del estimador es que el sesgo sea nulo y en ese caso diremos que el estimador es insesgado. En caso contrario el estimador sobrestima o infraestima al parámetro según que el sesgo sea positivo o negativo respectivamente.

11. ¿Qué significa error de tipo I?

Respuesta: Cuando del resultado de un contraste de hipótesis debemos rechazar la hipótesis nula, puede ocurrir que está fuese cierta, en cuyo caso habremos cometido un error que se denomina de tipo I.

12. Explique cuál es el efecto sobre la amplitud de un intervalo de confianza, con un nivel de confianza dado, de un aumento en el tamaño de la muestra aleatoria.

Respuesta: Al aumentar el tamaño de la muestra, disminuye la amplitud del intervalo.

13. Indicar la relación existente entre un contraste de hipótesis paramétrico y el error de tipo I.

Respuesta: Si θ es un parámetro desconocido y se desea contrastar la hipótesis nula H_o de que $\theta \in \Omega_o$ frente a la hipótesis alternativa H_1 de que $\theta \notin \Omega_o$, se denomina error de tipo I al que se comete cuando con el resultado del contraste, se rechaza la hipótesis nula cuando en realidad es cierta.

Si $\alpha(\theta) =$ Probabilidad del error de tipo I $= P(\text{rechazar } H_o \,/\, H_o \text{ es cierta})$, se llama tamaño del error de tipo I, o nivel de significación, a $\alpha = \max_{\theta \in \Omega_0} \alpha(\theta)$

Cuando la hipótesis nula H_o es simple $(\theta = \theta_o)$, la probabilidad de error de tipo I coincide con el nivel de significación.

14. ¿Existe alguna relación entre el error de tipo I y el error de tipo II?. Explique la respuesta.

Respuesta: Para un tamaño fijo de la muestra n, al aumentar el nivel de significación α, que es la probabilidad de rechazar H_o siendo falsa (error de tipo I), disminuye la probabilidad β de aceptar H_o siendo falsa (error de tipo II) y viceversa.
Si el tamaño de la muestra aumenta, es posible que α y β disminuyan, pero para una nivel de significación α fijo la probabilidad β de error de tipo II disminuye.

15. Explique conceptualmente porqué es importante que un estimador sea eficiente.

Respuesta: Un estimador es eficiente si es insesgado y su varianza alcanza la cota de Cramer-Rao, es decir, tiene menor varianza que cualquier otro estimador insesgado. Ello es importante porque, bajo la hipótesis de eficiencia, el estimador toma para diferentes muestras valores próximos unos a otros.

16. ¿Qué entendemos por error cuadrático medio?

Respuesta: Se define el error cuadrático medio de un estimador $\hat{\theta}$: $ECM(\hat{\theta}) = E\left[(\hat{\theta} - \theta)^2\right]$ donde θ es el parámetro a estimar. Si se desarrolla esta expresión se obtiene que:
$ECM(\hat{\theta}) = Var(\hat{\theta}) + \left(E(\hat{\theta}) - \theta\right)^2$,
denominándose sesgo de $\hat{\theta}$ a la diferencia $\left[E(\hat{\theta}) - \theta\right]$

17. ¿Cuál es el objetivo del contraste χ^2 de Pearson de bondad de ajuste y qué estadístico utiliza?.

Respuesta: Tiene por objetivo verificar si una muestra procede de una población con una determinada probabilidad. Se utiliza el estadístico $\chi^2 = \frac{\sum_{i=1}^{k} (n_i - n\,p_i)^2}{n\,p_i}$ que se distribuye aproximadamente como una χ^2 con (k-1) grados de libertad, siendo n el tamaño de la población en la que hemos efectuado una partición en 'k' clases, n_i el número de observaciones de la clase i-ésima, y p_i la probabilidad de que una observación caiga en la clase i-ésima.

18. Explique la diferencia entre estimador y estimación. Ponga un ejemplo.

Respuesta: En una población consideramos una muestra de tamaño n y sean $X_1, X_2, \cdots, X_n$ las n variables muestrales. Sea θ un parámetro muestral. Se llama estimador de θ a una cierta función de la muestra:

$\hat{\theta} = g(X_1, X_2, \cdots, X_n)$ elegida con ciertas propiedades de idoneidad.

Efectuada una realización muestral $(x_1, x_2, \cdots, x_n)$, se llama estimación, al valor del estimador para esa realización: $\hat{\theta} = g(x_1, x_2, \cdots, x_n)$

Ejemplo.- Supongamos una cierta población, una variable aleatoria X y que desconocemos la media poblacional μ, que deseamos estimar. Elegimos un tamaño muestral, por ejemplo, n = 5 y utilizamos como estimador la media muestral:

$$\hat{\mu} = \frac{\sum_{i=1}^{5} x_i}{5}$$

Sea por ejemplo (3, 5, 8, 3, 7) una realización muestral. Entonces $\frac{3+5+8+3+7}{5} = 5{,}2$ es una estimación de μ

19. Explique conceptualmente que mide la potencia de un contraste

Respuesta: La potencia de un contraste es la probabilidad de rechazar la hipótesis nula H_0. Si esta es cierta, la potencia del contraste coincide con el error de tipo I, y si H_0 es falsa, la potencia del contraste sería $(1-\beta)$, donde β es el error de tipo II, es decir, la probabilidad de aceptar H_0 siendo falsa.

20. Teorema de Chebychef, explique su significado.

Respuesta: Si X es una variable aleatoria, tal que $E(X) = \mu$ y $Var(X) = \sigma^2$, entonces $\forall k > 0$ se cumple:

$$P\left[\,|X - \mu| \geq k\right] \leq \frac{\sigma^2}{k^2}$$

Significa que la probabilidad de que un valor de la variable se aleje de la media de la población una distancia superior a k, está acotada por el cociente $\frac{\sigma^2}{k^2}$, es decir, es menor cuanto menor sea la varianza o cuanto mayor sea el cuadrado de k.

21. ¿Cuando se utiliza la desigualdad de Chebychef para obtener intervalos de confianza?. Razone la respuesta.

Respuesta: Cuando se desea un intervalo de confianza para la media y se desconoce la distribución de la población pero se conoce la varianza.

En efecto, para una muestra de tamaño n de una población con media μ desconocida y varianza σ^2 conocida. Utilizando como estimador de μ la media muestral $\overline{x}$, sabemos que:

$$E(\overline{x}) = \mu \qquad \mathrm{Var}(\overline{x}) = \frac{\sigma^2}{n}$$

Sustituyendo en la desigualdad de Chebychef $P\left[\left|\overline{x} - E(\overline{x})\right| \le k\right] \ge 1 - \frac{\mathrm{Var}(\overline{x})}{k^2}$, queda:

$P\left[\left|\overline{x} - \mu\right| \le k\right] \ge 1 - \frac{\sigma^2}{n k^2}$ que equivale $P\left[\overline{x} - k \le \mu \le \overline{x} + k\right] \ge 1 - \frac{\sigma^2}{n k^2}$

La expresión $1 - \frac{\sigma^2}{n k^2}$ es el nivel de confianza $(1-\alpha)$: $1 - \frac{\sigma^2}{n k^2} = 1 - \alpha$, de donde, $k = \frac{\sigma}{\sqrt{n\alpha}}$.

En definitiva, el intervalo de confianza buscado sería: $\left[\overline{x} - \frac{\sigma}{\sqrt{n\alpha}}\ ,\ \overline{x} + \frac{\sigma}{\sqrt{n\alpha}}\right]$

22. Explique razonadamente la diferencia entre un contraste de independencia y uno de homogeneidad.

Respuesta: En un contraste de independencia se establece la hipótesis nula H_o: X e Y son independientes, donde X e Y son dos características de la población.

Para realizar el contraste se utiliza el estadístico $\chi^2 = \sum_{i=1}^{k}\sum_{j=1}^{m} \frac{(O_{ij} - e_{ij})^2}{e_{ij}}$, donde O_{ij} son las frecuencias observadas de la muestra y e_{ij} son las frecuencias esperadas bajo el supuesto de independencia.

En un contraste de homogeneidad se parte de r muestras independientes clasificadas en s características. La hipótesis nula: $H_o: p_{1j} = p_{2j} = p_{3j} = \cdots\cdots = p_{rj}$ donde $(j = 1, 2, \cdots, s)$, es decir, la probabilidad de cada categoría es la misma en todas las muestras.

Para realizar el contraste se utiliza el mismo estadístico que el de independencia.

23. ¿Cuándo se puede decir que un estimador es insesgado?. Razone la respuesta poniendo algún ejemplo?

Respuesta: Un estimador $\hat{\theta}$ de un parámetro θ es insesgado cuando $E(\hat{\theta}) = \theta$

Por ejemplo, la media muestral $\overline{x}$ es insesgada para determinar la media poblacional:

$$E(\overline{x}) = E\left(\frac{\sum_{i=1}^{k} x_i}{k}\right) = \frac{\sum_{i=1}^{k} E(x_i)}{k} = \frac{k\,\mu}{k} = \mu$$

24. Estimador eficiente. Expresión analítica.

Respuesta: Un estimador $\hat{\theta}$ del parámetro θ es eficiente si es insesgado y de varianza mínima, es decir, coincide con la cota de Frechet-Cramer-Rao: $\text{Var}(\hat{\theta}) = \dfrac{1}{n E\left[\left(\dfrac{\vartheta \ln f(x\,;\theta)}{\vartheta\,\theta}\right)^2\right]}$

25. Relación existente entre estimador consistente y estimador consistente en media cuadrática.

Respuesta: Un estimador $\hat{\theta}_n$ de un parámetro θ, obtenido para una muestra de tamaño n, es consistente (consistencia simple o en probabilidad) si la sucesión $\left\{\hat{\theta}_n\right\}$ converge en probabilidad a θ, es decir:

$$\lim_{n \to \infty} P\left[\left|\hat{\theta}_n - \theta\right| < \varepsilon\right] = 1 \qquad \forall\, \varepsilon > 0$$

Un estimador es consistente en media cuadrática cuando $\lim\limits_{n \to \infty} E(\hat{\theta}_n - \theta)^2 = 0$

Si se verifica que un estimador es consistente en media cuadrática, también lo es en probabilidad. El reciproco no es necesariamente cierto.

26. Explique brevemente en que consiste el método de los momentos y cuales son las propiedades de los estimadores obtenidos por éste método.

Respuesta: Consiste en igualar tantos momentos muestrales como parámetros a estimar, a los correspondientes momentos poblacionales.
Si los parámetros a estimar son momentos poblacionales respecto del origen, los estimadores son insesgados. Pero, en general no lo son y por tanto no serían eficientes.
Bajo condiciones bastante generales, son consistentes.
Para estimar momentos poblacionales, son asintóticamente normales.

27. Indicar si los estimadores obtenidos por el método de los momentos son eficientes o no.

Respuesta: En general no son eficientes ya que no siempre el estimador obtenido por el método de los momentos es insesgado.

28. ¿Qué características pueden considerarse esenciales en el planteamiento de un contraste paramétrico?

Respuesta: Formulación de la hipótesis nula H_0 y de la hipótesis alternativa H_1 en términos estadísticos. Ambas deben de ser mutuamente excluyentes.
Determinación del test estadístico o estadístico de prueba apropiado.
Selección del nivel de significación α.
Determinación de la región crítica.
Selección aleatoria de la muestra.
Establecimiento de la regla de decisión y su interpretación.

29. ¿Cuándo se deberá utilizar un contraste de independencia y cuando uno de homogeneidad ?

Respuesta: Se utiliza un contraste de independencia cuando se trata de contrastar si existe dependencia entre dos características de la misma población.
Se utiliza un contraste de homogeneidad cuando se desea contrastar si dos o más muestras proceden de la misma población.

30. Explique conceptualmente para qué se puede utilizar la función de distribución empírica de la muestra.

Respuesta: La función de distribución empírica de la muestra $F_n(x)$ converge en probabilidad a la función de distribución de la población, al aumentar el tamaño de la muestra. Luego su gráfica puede utilizarse para determinar la forma general de la distribución poblacional.

31. ¿Qué entendemos por muestra aleatoria simple?

Respuesta: Dada una variable aleatoria simple X con función de distribución $F(x)$, se denomina muestra aleatoria simple de tamaño n al conjunto de n variables $(x_1, x_2, \cdots, x_n)$ independientes, cada una de ellas distribuida idénticamente igual que la variable X.

32. Clasifique los resultados posibles de la decisión tomada en un contraste de hipótesis utilizando la información proporcionada por una muestra, respecto de la naturaleza de la hipótesis nula. Razone la respuesta.

Respuesta: La hipótesis nula H_0 puede ser verdadera o falsa. Pueden presentarse los siguientes resultados:

	H_0 verdadera	H_0 falsa
Aceptamos H_0	Decisión correcta	Error de tipo II
Rechazamos H_0	Error de tipo I	Decisión correcta

33. ¿En qué contrastes podemos utilizar el estadístico χ^2 de Pearson?, ¿Para qué se utiliza?

Respuesta: En contrastes sobre la varianza σ^2 de una población normal $N(\mu, \sigma)$ con la media poblacional μ desconocida. El estadístico $\frac{(n-1)\,s^2}{\sigma^2} \approx \chi^2_{n-1}$ se distribuye como una χ^2 con (n-1) grados de libertad.
Se utiliza para determinar la región crítica (y de aceptación) para un nivel de confianza α dado.

Por ejemplo, si el contraste es:

$$H_0 : \sigma^2 = \sigma_0^2$$
$$H_1 : \sigma^2 \neq \sigma_0^2$$

entonces la región crítica viene definida por las desigualdades: $\frac{(n-1)\,s^2}{\sigma^2} < \chi^2_{\alpha/2\,;\,(n-1)}$ o $\frac{(n-1)\,s^2}{\sigma^2} > \chi^2_{1-\alpha/2\,;\,(n-1)}$

34. ¿Qué es una hipótesis estadística?, ¿Qué es una hipótesis nula?. Razone la respuesta.

Respuesta: Una hipótesis estadística es una afirmación verdadera o falsa acerca del valor de alguna característica desconocida de la población.
Para efectuar un contraste de hipótesis, aceptamos una hipótesis como verdadera, a la cual denominamos hipótesis nula, frente a otra complementaria que llamamos hipótesis alternativa.

35. ¿Cual es el objetivo de los contrastes de aleatoriedad?. Razone la respuesta.

Respuesta: Tienen por objetivo determinar si la muestra elegida en el proceso de muestreo es aleatoria.

36. Definición de estadístico, estimación y estimador. Ponga un ejemplo de cada uno de estos conceptos en los que se aprecien las diferencias entre ellos.

Respuesta:
Estadístico es cualquier función de la variable muestral $(x_1, x_2, \cdots, x_n)$: $\hat{\theta} = g(x_1, x_2, \cdots, x_n)$

Estimador es un estadístico que usamos para hallar el valor de un parámetro poblacional desconocido.

Estimación es un valor concreto del estimador, para una muestra concreta elegida.

Por ejemplo: Estadísticos serían el máximo de $(x_1, x_2, \cdots, x_n)$; la media muestral $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$; la cuasivarianza muestral $s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$; o cualquier momento muestral respecto al origen

$\alpha_n = \frac{1}{n}\sum_{i=1}^{n} x_i^n$, etc.

Un estimador de la media poblacional sería, por ejemplo, la media muestral $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

Si como estimador de la media poblacional se elige la media muestral, para muestras de tamaño 5, elegida la muestra (2,5, 3, 2, 3,1 , 2,7), la estimación correspondiente sería:

$$\frac{2{,}5 + 3 + 2 + 3{,}1 + 2{,}7}{5} = 2{,}66$$

37. Concepto de error cuadrático medio. ¿Para qué se utiliza?

Respuesta: Sea $\hat{\theta}$ el estimador de un parámetro poblacional θ, el error cuadrático medio se define como el valor de $ECM(\hat{\theta}) = E\left[(\hat{\theta} - \theta)^2\right]$.

Desarrollando la expresión se obtiene: $ECM(\hat{\theta}) = Var(\hat{\theta}) + \left(E(\hat{\theta}) - \theta\right)^2$, denominándose sesgo de $\hat{\theta}$ a la diferencia $\left(E(\hat{\theta}) - \theta\right)$. En efecto:

$ECM(\hat{\theta}) = E\left[(\hat{\theta} - \theta)^2\right]$ = (restamos y sumamos $E(\hat{\theta})$) = $E\left[(\hat{\theta} - E(\hat{\theta}) + E(\hat{\theta}) - \theta)^2\right] =$

$$= E\left[(\hat{\theta} - E(\hat{\theta}))^2\right] + E\left[(E(\hat{\theta}) - \theta)^2\right] + 2E\left[(\hat{\theta} - E(\hat{\theta}))\ (E(\hat{\theta}) - \theta)\right] =$$

$= Var(\hat{\theta}) + (E(\hat{\theta}) - \theta)^2 + 2\left[(E(\hat{\theta}) - \theta)\right] E\left[(\hat{\theta} - E(\hat{\theta}))\right]$ = teniendo en cuenta que

$$E\left[(\hat{\theta} - E(\hat{\theta}))\right] = 0 = Var(\hat{\theta}) + (E(\hat{\theta}) - \theta)^2 = Var(\hat{\theta}) + (sesgo)^2$$

En el problema de la estimación puntual que el error cuadrático medio sea lo menor posible, lo cual se consigue cuanto menores sean la varianza de $\hat{\theta}$ y el valor absoluto del sesgo. Cuando el estimador es insesgado [$sesgo(\hat{\theta}) = 0$] el ECM coincide con la varianza $ECM(\hat{\theta}) = Var\left[(\hat{\theta})\right]$

38. Relacione las propiedades de los estimadores de máxima verosimilitud.

Respuesta:

Son consistentes.
En general, son insesgados. En caso contrario, son insesgados asintóticamente.
Si $\hat{\theta}$ es eficiente, entonces es máximo verosímil y único. Aunque no todo estimador máximo verosímil es eficiente, aunque si lo es asintóticamente.
Si $\hat{\theta}$ es máximo verosímil entonces es asintóticamente normal $N\left(\hat{\theta}, \sqrt{Var(\hat{\theta})}\right)$, en donde $Var(\hat{\theta})$ coincide con la cota de Frechet-Cramer-Rao.
Si $\hat{\theta}$ es eficiente, entonces el estimador máximo verosímil de θ, si es único, es función de $\hat{\theta}$.
Principio de invarianza de Zehna.- Si $\hat{\theta}$ es máximo verosímil, es invariante por una transformación biunívoca.

39. Condiciones que debe cumplir el pivote en el método pivotal.

Respuesta: Si θ es el parámetro desconocido, entonces el pivote es una función $T(X_1, X_2, \cdots, X_n, \theta)$ que debe cumplir las condiciones:

El valor que toma $T(X_1, X_2, \cdots, X_n, \theta)$ para cada muestra depende exclusivamente de θ
La distribución muestral del estadístico $T(X_1, X_2, \cdots, X_n, \theta)$ no depende de θ

40. Concepto de nivel de significación y potencia de un contraste. Relación entre ambos.

Respuesta: El nivel de significación α de un contraste es la probabilidad de cometer error de tipo I, es decir, es la probabilidad de rechazar la hipótesis nula, siendo ésta cierta. También se denomina tamaño de la región crítica (o de rechazo) ya que la probabilidad de que el estimador pertenezca a la región crítica es precisamente α.
La potencia del contraste es la probabilidad de rechazar la hipótesis nula. Si denominamos por β la probabilidad de error de tipo II, es decir, la probabilidad de aceptar la hipótesis nula siendo falsa, entonces:

$$\text{Potencia del contraste} = \begin{cases} \alpha & \text{si } H_0 \text{ es cierta} \\ 1-\beta & \text{si } H_0 \text{ es falsa} \end{cases}$$

Para un tamaño muestral n fijo, si α aumenta, entonces β disminuye y, por lo tanto, $1-\beta$ aumenta y viceversa.

41. ¿Cual es la interpretación del concepto 'grados de libertad' a la hora de utilizar estimadores?

Respuesta: En una muestra de tamaño n, las n variables $(X_1, X_2, \cdots, X_n)$ son independientes. Se suele decir que el conjunto de las n variables $(X_1, X_2, \cdots, X_n)$ contiene n grados de libertad. Ahora bien, es posible que un estimador cualquiera $\hat{\theta} = \hat{\theta}(X_1, X_2, \cdots, X_n)$ mantenga los n grados de libertad o no.

Por ejemplo, en una población en la que desconocemos la media poblacional, utilizamos como estimador
$\hat{\theta} = \sum_{i=1}^{n} (x_i - \overline{x})^2$ (para hacer estimaciones sobre la varianza) ocurre que hemos perdido un grado de libertad, puesto que sabemos que $\sum_{i=1}^{n} x_i = n\,\overline{x}$, de donde, conociendo solamente (n-1) valores de la muestra podemos despejar el que queda. Así, pues, $\hat{\theta}$ posee (n-1) grados de libertad.

42. Teorema de factorización de Fisher-Neyman

Respuesta:

Si $(X_1, X_2, \cdots, X_n)$ es una muestra aleatoria simple, siendo $P(x_1, x_2, \cdots, x_n, \theta)$ ó $f(x_1, x_2, \cdots, x_n, \theta)$ las funciones de probabilidad o densidad, respectivamente, entonces el estadístico

$T = T(X_1, X_2, \cdots, X_n)$ es suficiente para el parámetro θ si y sólo si existen dos funciones $g\left[T(X_1, X_2, \cdots, X_n), \theta\right]$ y $h(x_1, x_2, \cdots, x_n)$, tales que:

$$P(x_1, x_2, \cdots, x_n, \theta) = g\left[T(X_1, X_2, \cdots, X_n), \theta\right] \; h(x_1, x_2, \cdots, x_n)$$

ó

$$f(x_1, x_2, \cdots, x_n, \theta) = g\left[T(X_1, X_2, \cdots, X_n), \theta\right] \; h(x_1, x_2, \cdots, x_n)$$

43. Lema de Neyman-Pearson

Respuesta:

Proporciona un criterio para hallar la región crítica de tamaño α en un contraste de hipótesis que haga mínimo el error de tipo II, $\beta = P(\text{Aceptar } H_o \,/\, H_o \text{ es falsa})$. Diremos que es la mejor región crítica.

Establece que, obtenida una muestra aleatoria simple $(X_1, X_2, \cdots, X_n)$ de una población con función de densidad $f(x\,;\theta)$, siendo $L(x_1, x_2, \cdots, x_n\,;\theta) = \prod_{i=1}^{n} f(x_i\,;\theta)$ la función de verosimilitud de una muestra, si queremos efectuar el contraste:

$$\begin{cases} H_o: \theta = \theta_o \\ H_1: \theta = \theta_1 \end{cases}$$

sea $k > 0$ y C un subconjunto del espacio muestral R^n tal que:

$$\begin{cases} \dfrac{L(x_1, x_2, \cdots, x_n\,;\theta_o)}{L(x_1, x_2, \cdots, x_n\,;\theta_1)} \le k \quad \text{si} \quad (x_1, x_2, \cdots, x_n) \in C \\[2ex] \dfrac{L(x_1, x_2, \cdots, x_n\,;\theta_o)}{L(x_1, x_2, \cdots, x_n\,;\theta_1)} > k \quad \text{si} \quad (x_1, x_2, \cdots, x_n) \notin C \end{cases}$$

Entonces C es la mejor región crítica de tamaño α para efectuar el contraste.

44. Intervalo de confianza de una proporción para muestras pequeñas.

Respuesta: Consideramos el estimador de máxima verosimilitud que, para una muestra de tamaño n de una población Bernouilli $B(1, p)$, es la proporción muestral $\hat{p} = \dfrac{n^o \text{ éxitos en n pruebas}}{n^o \text{ de pruebas}} = \dfrac{X}{n}$ cuya función de probabilidad $P\left[\dfrac{X}{n} = x\right] = P(X = n\,x) = \dbinom{n}{n\,x} p^{n\,x}\,(1-p)^{n-n\,x}$, donde $x = 0, \dfrac{1}{n}, \dfrac{2}{n}, \; \cdots \; , 1$

Los extremos p_i y p_s de un intervalo de nivel de confianza $(1-\alpha)$ el 100%, verificarían las ecuaciones:

$$P\left[\frac{X}{n} \le p_i\right] = \sum_{x=0}^{p_i} \binom{n}{n\,x} p^{n\,x}\,(1-p)^{n-n\,x} \le \frac{\alpha}{2}$$

$$P\left[\frac{X}{n} \ge p_s\right] = \sum_{x=p_s}^{1} \binom{n}{n\,x} p^{n\,x}\,(1-p)^{n-n\,x} \le \frac{\alpha}{2}$$

dada la dificultad de resolución de estas ecuaciones, Clopper y Pearson elaboraron unos ábacos que proporcionan las soluciones para distintos valores de x, con n y α fijos.

45. Contrastes uniformemente más potentes.

Respuesta: Supongamos un contraste de una hipótesis simple H_o frente a una hipótesis compuesta H_1, se dice que C es *la región crítica uniformemente más potente de tamaño* α si es la mejor región crítica de ese tamaño para contrastar H_o, para cada una de las hipótesis simples de las que consta H_1.
Si la región crítica de un contraste cumple esta propiedad, decimos que el contraste es el uniformemente más potente de tamaño α.

46. Regla de decisión para el contraste de Pearson de bondad de ajuste cuando los parámetros poblacionales son conocidos.

Respuesta: Consideremos una partición en k clases del campo de variación de la variable aleatoria X y una muestra aleatoria simple de tamaño n. Sea n_i, $(i = 1, 2, \cdots, k)$, el número de observaciones muestrales contenidas en la clase i-ésima.
Supongamos que la hipótesis nula H_o establece las probabilidades $p_1, p_2, \cdots, p_k$ de que una observación caiga en cada una de las clases. Entonces, rechazaremos H_o, al nivel de significación α, si:

$$\chi^2_{exp} = \frac{\sum_{i=1}^{k} (n_i - n\,p_i)^2}{n\,p_i} > \chi^2_{\alpha\,;(k-1)}$$

donde $\chi^2_{1-\alpha}$ es tal que

$$P\left[\chi^2_{k-1} > \chi^2_{1-\alpha} \;/\; H_o\right] = \alpha$$

siendo χ^2_{k-1} una chi-cuadrado con (n-1) grados de libertad.

47. ¿Cuándo decimos que un estimador es UMVUE?

Respuesta: Un estimador $\hat{\theta}_o$ es UMVUE (insesgado y uniformemente de mínima varianza) para estimar el parámetro θ, si dado cualquier otro estimador $\hat{\theta}$ de θ, se verifica que $Var(\hat{\theta}_o) \leq Var(\hat{\theta})$, para todos los posibles valores de θ.

48. ¿De qué depende que la amplitud de un intervalo de confianza para la media, siendo la varianza desconocida, sea mayor o menor?

Respuesta:

Para un tamaño de la muestra fijo, a mayor nivel de confianza $(1-\alpha)$, mayor amplitud del intervalo. Para un nivel de confianza fijo $(1-\alpha)$, a mayor tamaño de la muestra, menor amplitud del intervalo.

49. Discutir la siguiente aseveración <<Los estimadores insesgados siempre dan mejores estimaciones que los estimadores sesgados>>.

Respuesta: No es cierto en general. Entre dos estimadores consideraremos mejor el que proporciona un menor error cuadrático medio $ECM(\hat{\theta}) = Var(\hat{\theta}) + (sesgo(\hat{\theta}))^2$. Dados dos estimadores, $\hat{\theta}_1$ sesgado y $\hat{\theta}_2$ insesgado, podría ocurrir que $Var(\hat{\theta}_1) + (sesgo(\hat{\theta}_1))^2 < Var(\hat{\theta}_2)$

50. ¿Cual es la utilidad del lema de Neyman-Pearson?

Respuesta: Proporciona un criterio para hallar la región crítica de tamaño α en un contraste de hipótesis, que haga mínimo el error de tipo II, $\beta = P(\text{Aceptar } H_o \ / \ H_o \text{ es falsa})$.

51. Intervalo de confianza de una proporción para muestras grandes.

Respuesta: Consideramos una población con una variable que se distribuya $B(1, p)$ (distribución de Bernouilli de parámetro p). Sea $\hat{p}$ la proporción muestral de éxitos para una muestra de tamaño n suficientemente grande, y $\hat{q} = 1 - \hat{p}$.

Sea Z normal $N(0,1)$ y $z_{\frac{\alpha}{2}}$ tal que $P\left[Z \geq z_{\frac{\alpha}{2}}\right] = \frac{\alpha}{2}$

Un intervalo de confianza, a un nivel $(1-\alpha)$, para la proporción poblacional p es:

$$I_{(1-\alpha)} = \left[\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\,\hat{q}}{n}} \ , \ \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\,\hat{q}}{n}}\right]$$

52. Tamaño de la muestra para estimar la proporción de una población.

Respuesta: Se trata de hallar el tamaño muestral n para construir un intervalo de confianza de longitud L, al nivel de confianza $(1-\alpha)$.

Como el intervalo de confianza es $I_{(1-\alpha)} = \left[\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\,\hat{q}}{n}} \ , \ \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\,\hat{q}}{n}}\right]$

cuya longitud es $L = 2\, z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\,\hat{q}}{n}}$, despejando se obtiene $n = \dfrac{4\, z^2_{\frac{\alpha}{2}}\, \hat{p}\,\hat{q}}{L^2}$

En particular, como el error $\varepsilon = \left|\hat{p} - p\right| = \dfrac{L}{2}$, se tiene que $n = \dfrac{z^2_{\frac{\alpha}{2}}\, \hat{p}\,\hat{q}}{\varepsilon^2}$

53. Los sucesos A y B verifican P(A)=0,4, P(B)=0,3 y $P(A\cap B)=0,1$. La probabilidad de $P(A/A\cup B)$:

a) 4/7 b) 5/6
c) 2/3 d) Ninguna respuesta es correcta

Respuesta: (c)

$$P(A/A\cup B)=\frac{P\left[A\cap(A\cup B)\right]}{P(A\cup B)}=\frac{P\left[(A\cap A)\cup(A\cap B)\right]}{P(A\cup B)}=\frac{P\left[(A\cup(A\cap B)\right]}{P(A\cup B)}=\frac{P(A)}{P(A\cup B)}=\frac{0,4}{0,4+0,3-0,1}=\frac{0,4}{0,6}=\frac{2}{3}$$

54. Sea A un suceso contenido en el suceso B, $A\subset B$. ¿Qué opción es correcta?.

a) $P(B)<P(A)$ b) $P(B-A)=P(B)-P(A)$
c) $P(A\cap B)=0$ d) Ninguna de las respuestas

Respuesta: (b)

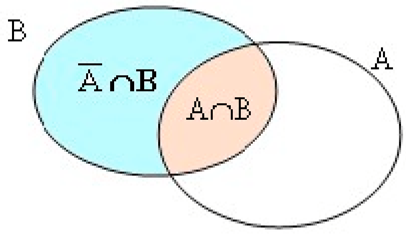
Si $A\subset B \mapsto B=A\cup(B-A)$, como A y $B-A$ son incompatibles: $P(B)=P(A)+P(B-A)$
con lo cual, $P(B-A)=P(B)-P(A)$

55. Sean A y B dos sucesos cualesquiera con $P(B)>0$, se verifica:

a) $P(A/B)=P(\overline{A}/B)$ b) $P(A\cap B)<P(A)$
c) $P(A/B)+P(B/A)=1$ d) $P(A/B)+P(\overline{A}/B)=1$

Respuesta: (d)

Los sucesos $A\cap B$ y $\overline{A}\cap B$ constituyen una partición disjunta de B.



Luego $P(B)=P(A\cap B)+P(\overline{A}\cap B)$

de donde, se deduce que:

$$P(A/B)+P(\overline{A}/B)=\frac{P(A\cap B)}{P(B)}+\frac{P(\overline{A}\cap B)}{P(B)}=\frac{P(A\cap B)+P(\overline{A}\cap B)}{P(B)}=\frac{P(B)}{P(B)}=1$$

56. Una fabrica de perfumes tiene dos marcas en el mercado (A y B), y proyecta el lanzamiento de una nueva marca. La probabilidad de compra de A es 0,3, la de B es 0,5, y la de A y B es 0,1.

Antes de lanzar la nueva marca, la fabrica necesita conocer la probabilidad de que no se compre ni A ni B, así como la probabilidad de que solo se compre una de las dos marcas. Señalar la opción correcta:

a) 0,2 y 0,6 b) 0,3 y 0,6
c) 0,8 y 0,6 d) Ninguna respuesta es correcta

Respuesta: (b)

El suceso no comprar ni A ni B: $\overline{A}\cap\overline{B}=(A\cup B)^c$, y como son sucesos disjuntos,
$(A\cup B)\cup(A\cup B)^c=\Omega$

$$P(A\cup B)^c=1-P(A\cup B)=1-\left[P(A)+P(B)-P(A\cap B\right]=1-(0,3+0,5-0,1)=0,3$$

con lo cual, $P(\overline{A}\cap\overline{B})=P(A\cup B)^c=0,3$

Comprar la marca A y no comprar la B es (A-B). Análogamente, comprar la marca B y no comprar la A es (B-A). En consecuencia, la probabilidad de comprar una de las marcas será:

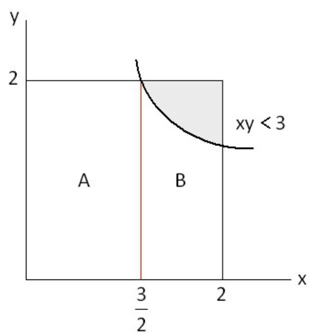
$$\text{P(una de las marcas)}=P\left[(A-B)\cup(B-A)\right]=P(A-B)+P(B-A)=\left[P(A)-P(A\cap B\right]+\left[P(B)-P(A\cap B\right]=$$

$$=P(A)+P(B)-2.P(A\cap B)=0,3+0,5-0,2=0,6$$

57. Se eligen al azar dos puntos en el intervalo (0,2). Señalar la probabilidad de que el producto de ambos sea menor que 3.

a) 0,75 b) 1
c) 0,92 d) Ninguna respuesta es correcta

Respuesta: (d)



El suceso $xy<3$ está formado por los puntos del cuadrado de lado 2 que se encuentran por debajo de la hipérbola $xy=3$, es decir, pertenecientes a los recintos A o B.

La probabilidad de que se verifique $xy<3$ es igual a la de que un punto w pertenezca a uno de los recintos A o B.

$$P(xy<3)=P\left[(w\in A)\cup(w\in B)\right]=\frac{\text{Área A}+\text{Área B}}{\text{Área cuadrado}}=\frac{2.\frac{3}{2}+\int_{3/2}^{2}\frac{3}{x}\,dx}{4}=\frac{3}{4}+\frac{3}{4}.(\ln 2-\ln\frac{3}{2})=0,966$$

58.- Sean A y B dos sucesos correspondientes a un experimento aleatorio, tales que $A\cup B=\Omega$ con $P(A)=0,8$ y $P(B)=0,5$. Se pide calcular las probabilidades:

<1> $P(A\cap B)$ <2> $P(A\cup\overline{B})$

<3> $P(\overline{A}\cup B)$ <4> $P(\overline{A}\cup\overline{B})$

Respuesta:

Sabemos que $P(A\cup B)=P(A)+P(B)-P(A\cap B)$, con la hipótesis $A\cup B=\Omega$, siendo $P(\Omega)=1$, se tiene: $1=0,8+0,5-P(A\cap B) \quad\Rightarrow\quad P(A\cap B)=0,3$

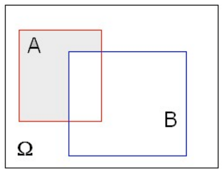
$\overline{B}$ es el complementario de B, en consecuencia, $\overline{B}=\Omega-B$

$$P(A\cup\overline{B})=P(A)+P(\overline{B})-P(A\cap\overline{B})$$

$$P(\overline{B})=P(\Omega)-P(B) \quad\mapsto\quad P(\overline{B})=1-0,5=0,5$$

de otra parte,

$$\overline{B}=\Omega-B=(A\cup B)-B=A\cap\overline{B} \quad\mapsto\quad P(\overline{B})=P(A\cap\overline{B})=0,5$$



con lo que, $P(A\cup\overline{B})=0,8+0,5-0,5=0,8$

también podríamos haber considerado: $A\cup B=\Omega \quad\Rightarrow\quad B\subset A$

con lo cual, $P(A\cup\overline{B})=P(A)=0,8$

$$P(\overline{A}\cup B)=P(B)=0,5$$

$$P(\overline{A}\cup\overline{B}) \overbrace{=}^{\text{Leyes de Morgan}} P(\overline{A\cap B})=1-P(A\cap B)=1-O,3=0,7$$