INGENIERÍA INFORMÁTICA DE GESTIÓN
Septiembre 2005
SOLUCION

1.- Sea ($X_1,........,X_n$) una muestra aleatoria simple de una variable aleatoria X con E(X) = μ y Var (X) = $k \cdot \mu^2$. Considérense los siguientes estimadores del parámetro μ:

$$T_1 = X_1 \qquad\qquad T_2 = \frac{3X_1 - 2X_2 + X_3}{6}$$

a) ¿Son insesgados ambos estimadores? (1 punto)

$$E[T_1] = E[X_1] = \mu$$

$$E[T_2] = E[\frac{3X_1 - 2X_2 + X_3}{6}] = \frac{1}{6}3E[X_1] - \frac{2}{6}E[X_2] + \frac{1}{6}E[X_3] =$$

$$\frac{1}{2}\mu - \frac{1}{3}\mu + \frac{1}{6}\mu = \frac{1}{3}\mu$$

$$Sesgo[T_2] = \frac{2}{3}\mu$$

b) Calcular el error cuadrático medio de los dos estimadores. (1 punto)

$$E.C.M[T_1] = V[T_1] = V[X_1] = k \cdot \mu^2$$

$$E.C.M[T_2] = V[T_2] + sesgo[T_2]^2$$

$$V[T_2] = \frac{1}{6^2}[3^2 k \cdot \mu^2 + (-2)^2 k \cdot \mu^2 + k \cdot \mu^2] = \frac{7}{18}k \cdot \mu^2$$

$$E.C.M[T_2] = \frac{7}{18}k \cdot \mu^2 + (\frac{2}{3})^2 \mu^2 = \mu^2(\frac{7}{18}k + \frac{4}{9})$$

c) ¿Para qué valores de k es el estimador T_2 mejor que T_1 de acuerdo al criterio del E.C.M. ? (0.5 punto)

$$\mu^2(\frac{7}{18}k + \frac{4}{9}) \prec k \cdot \mu^2 \Rightarrow (\frac{7}{18}k + \frac{4}{9}) \prec k \Rightarrow \frac{8}{11} \prec k$$

2.- X es una variable aleatoria con función de densidad, $f(x) = \begin{cases} k\cdot x - 7 & \text{sí } 7 \le x < 8 \\ -k\cdot x + 9 & \text{sí } 8 \le x \le 9 \end{cases}$

a) Calcular el valor de k. (0.5 puntos)

$$1 = \int_7^8 (kx-7)dx + \int_8^9 (-kx+9)dx = [k\frac{x^2}{2} - 7x]_7^8 + [-k\frac{x^2}{2} + 9x]_8^9 = -k+2 \Rightarrow k = 1$$

b) Obtener la esperanza y la varianza de X. (0.5 puntos)

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx = \int_7^8 (x^2-7x)dx + \int_8^9 (-x^2+9x)dx = [\frac{x^3}{3} - 7\frac{x^2}{2}]_7^8 + [-\frac{x^3}{3} + 9\frac{x^2}{2}]_8^9 = 8$$

$$E[X^2] = \int_7^8 (x^3-7x^2)dx + \int_8^9 (-x^3+9x^2)dx = 64.16$$

$$V[X] = E[X^2] - E[X]^2 = 64.16 - 64 = 0.16$$

c) Hallar la función de distribución. (0.75 puntos)

$$F(x) = \int_{-\infty}^{x} f(z)dz$$

$$x < 7 \Rightarrow F(x) = 0$$

$$7 \le x < 8 \Rightarrow F(x) = \int_{-\infty}^{7} f(z)dz + \int_7^x f(z)dz = 0 + \int_7^x (z-7)dz = \frac{x^2}{2} - 7x + \frac{49}{2}$$

$$8 \le x \le 9 \Rightarrow F(x) = \int_{-\infty}^{7} f(z)dz + \int_7^8 f(z)dz + \int_8^x f(z)dz = 0 + \int_7^8 (z-7)dz + \int_8^x (-z+9)dz =$$

$$(\frac{z^2}{2} - 7z)_7^8 + (\frac{-z^2}{2} + 9z)_8^x = \frac{-x^2}{2} + 9x - \frac{79}{2}$$

$$x > 9 \Rightarrow F(x) = \int_{-\infty}^{7} f(z)dz + \int_7^8 f(z)dz + \int_8^9 f(z)dz + \int_9^x f(z)dz = 0 + \int_7^8 (z-7)dz + \int_8^9 (-z+9)dz + 0 =$$

$$(\frac{z^2}{2} - 7z)_7^8 + (\frac{-z^2}{2} + 9z)_8^9 = 1$$

d) Si se define la variable Y = 0.5X + 0.375 . ¿Probabilidad de que Y sea superior a 4? (0.75 puntos)

$$P(Y \succ 4) = P(0.5X + 0.375 \succ 4) = P(X > 7.25) = 1 - P(X < 7.25) = 1 - \int_{7}^{7.25} (x-7)dx = 1 - 0.0312 = 0.9688$$

3.- Supongamos que el tiempo en horas dedicado por los estudiantes a preparar el examen final de Estadística sigue una distribución normal. Se toma una muestra aleatoria de 10 estudiantes obteniéndose los siguientes resultados: Media de la muestra 15 horas y Cuasidesviación típica de la muestra 4 horas.

Sea X: Tiempo de estudio en horas para el examen de Estadística, X→ N(μ,σ)

a) Calcular dos estimadores insesgados para el tiempo medio y varianza del tiempo que los estudiantes dedican a preparar el examen final de Estadística. (0.5 puntos)

El estimador insesgado propuesto para μ es la media muestral puesto que $E(\bar{x}) = \mu$
El estimador insesgado propuesto para σ^2 es la cuasivarianza muestral puesto que $E(\hat{S}^2) = \sigma^2$

b) Construir intervalos de confianza para el tiempo medio y varianza del tiempo que los estudiantes dedican a preparar el examen final, a un 95% de confianza. (0.5 puntos)

$\mu \in \left[\bar{x} \pm t_{\alpha/2} \frac{\hat{S}}{\sqrt{n}} \right]$=[12,7898;17,21018] siendo $t_{14,0,975}$=2,14

$\sigma^2 \in \left[\frac{(n-1)\hat{S}^2}{\chi^2_{14,0,025}} ; \frac{(n-1)\hat{S}^2}{\chi^2_{14,0,975}} \right]$ = [39,7868;8,58237] siendo $\chi^2_{14,0,025}$ =5,63 y $\chi^2_{14,0,975}$ =26,1

c) Un profesor de la asignatura asegura que el tiempo medio que dedican sus alumnos a la preparación del examen es de 16 horas. ¿Se puede aceptar este supuesto a un 5% de significación?. Responder basándose en el apartado anterior. (0.75 puntos)

H_o: μ=16
H_1: μ≠16

Por dualidad entre intervalos de confianza y contraste de hipótesis, y debido a que el contraste de hipótesis es bilateral podemos aceptar la hipótesis del profesor a un 5% de significación puesto que μ=16 está incluido en el intervalo de confianza calculado en el apartado anterior.

d) Otro profesor que imparte la asignatura de Estructura de Datos afirma que el tiempo medio de estudio que emplean sus estudiantes es mayor que el tiempo medio empleado en la asignatura de Estadística. Para contrastar esta hipótesis se han tomado dos muestras aleatorias de 100 estudiantes de cada asignatura obteniéndose los siguientes resultados:

Tiempo medio de estudio de los 100 estudiantes de Estructura=17 horas.
Desviación típica del tiempo de los 100 estudiantes de Estructura =5 horas.
Tiempo medio de estudio de los 100 estudiantes de Estadística=14 horas.
Desviación típica del tiempo de los 100 estudiantes de Estadística =3 horas.
¿Hay evidencia suficiente para aceptar la hipótesis del profesor de Estructura de datos? (0.75 puntos)

H_o: $\mu_{Estruct} \geq \mu_{Estadist}$
H_1: $\mu_{Estruct} < \mu_{Estadist}$

Para discutirlo vamos a calcular el p-valor.

Estadístico de contraste $z = \dfrac{\bar{x} - \bar{y} - (\mu_x - \mu_y)}{\sqrt{\dfrac{s^2{}_x}{n} + \dfrac{S^2{}_y}{n}}} = \dfrac{17 - 14 - 0}{\sqrt{\dfrac{25}{100} + \dfrac{9}{100}}} = 5{,}14495$

Al ser el tamaño muestral suficientemente grande (n=100) podemos utilizar las varianzas muestrales como buenas aproximaciones de las varianzas poblacionales.

P-valor= P(z< 5,14495)=1
Podemos aceptar la hipótesis nula planteada.

4. La probabilidad de que una consola portátil de reciente aparición en el mercado tenga un defecto de fábrica es del 5%. Cada día se producen de forma independiente 60 consolas.

a) ¿Cuál es el numero medio y la desviación típica de consolas defectuosas que se producen en un día? (0.75 puntos)

Sea X una variable aleatoria que cuenta el número de consolas defectuosas que aparecen entre las n=60 consolas se producen un día, con una probabilidad de p=0,05 de ser defectuosa, entonces X~B(60; 0,05)

E(X) = np = 60×0,05 = 3

V(X) = np(1-p) = 60×0,05×(1-0,05) = 2,85 → DT(X) = √V(X) = √2,85 = 1,688

b) ¿Cuál es la probabilidad de que en un día no se produzca ninguna consola defectuosa? (0.5 puntos)

Del apartado anterior sea $X \sim B(60; 0{,}05)$

$P(X=0) = \binom{60}{0} (0{,}05)^0 (1-0{,}05)^{60} = 0{,}046$

c) Una cadena de electrodomésticos que compra las consolas directamente a la fábrica han comprobado que el número de consolas vendidas en un día sigue una distribución de Poisson, con una medía de 1 consola vendida por día. ¿Cuál es el número medio y la desviación típica de número de consolas vendidas por dicha cadena en una semana? (0.75 puntos)

Sea Y el número de consolas vendidas en un día sigue una distribución de Poisson, con una medía de 1 consola vendida por día, entonces $Y \sim P(1)$

Si Z es el número de consolas vendidas en una semana, entonces $Z = \sum_{i=1}^{7} Y_i$, entonces $E(Z) = E(\sum_{i=1}^{7} Y_i) = \sum_{i=1}^{7} E(Y_i) = 7$.

Así si el número medio de consolas vendidas en una semana $Z \sim P(7)$, entonces $V(Z) = 7 \rightarrow DT(Z) = \sqrt{V(Z)} = \sqrt{7} = 2{,}65$

d) ¿Qué probabilidad hay de que la cadena de electrodomésticos venda más de tres consolas en una semana (7 días)? (0.5 puntos)

Del apartado anterior, el número de consolas vendidas en una semana $Z \sim P(7)$

$P(Z>3) = 1-P(Z\leq 3) = 1-\sum_{j=0}^{3} P(Z = j) = 1-(0{,}0009+0{,}0064+0{,}0223+0{,}0521) = 1-0{,}0817 = 0{,}9183$