



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
FACULTAD DE MATEMÁTICAS / DEPARTAMENTO DE ESTADÍSTICA
EYP2405 Métodos Estadísticos

Guía 1

Profesora : Lorena Correa
Ayudante : Catalina Vallejos
Segundo Semestre del 2009

Propiedades de los Estimadores

1. El número de huevos puestos por un insecto sigue una distribución de Poisson con media λ desconocida. Cada uno de estos tiene una probabilidad p desconocida de originar una larva, comportándose los huevos de manera independiente. Un entomólogo estudia conjuntos de n de tales insectos observando el número de huevos puestos y si nacen larvas o no. Identifique la ley de probabilidad de los datos y el espacio paramétrico.
2. Sean X e Y variables aleatorias positivas con densidad:

$$f_X(x) = \frac{\lambda(\lambda x)^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} \quad \lambda > 0, \alpha > 0$$

$$f_Y(y) = \frac{y^{\alpha-1} e^{-y/\theta}}{\Gamma(\alpha)\theta^\alpha} \quad \theta > 0, \alpha > 0.$$

Verifique que estos modelos son efectivamente reparametrizaciones uno del otro.

3. Encuentre estadígrafos suficientes para estimar θ en base a una muestra de las siguientes distribuciones con a conocido :
 - a) $p(x, \theta) = \theta x^{\theta-1}, \quad 0 < x < 1, \theta > 0.$
 - b) $p(x, \theta) = \theta a x^{a-1} \exp\{-\theta x^a\}, \quad x > 0, \theta > 0, a > 0.$
 - c) $p(x, \theta) = \frac{\theta a^\theta}{x^{\theta+1}}, \quad x > a, \theta > 0, a > 0,$
4. Sean $X_1, X_2 \sim i.i.d. P(\lambda)$. Se desea estimar $P(X_1 = 0) = e^{-\lambda}$. Encuentre el E.C.M. de los siguientes estimadores:
 - a.- $\frac{1}{2} X_1 + X_2$
 - b.- Proporción de $X_i = 0$ (puede ser 0, 1/2 ó 1)
5. Sean X_1, X_2, \dots, X_n iid con densidad $\lambda e^{-\lambda x}, x > 0, n > 2$. Sea $S_n = \sum_{i=1}^n X_i$. Es bien conocido que $Z = \lambda S_n$ tiene densidad

$$f_Z(z) = z^{n-1} e^{-z} / (n-1)! \quad z \geq 0$$

Utilice esto para calcular el sesgo y el E.C.M. de $\hat{\lambda} = (n-1)/S_n$.

6. Sean $X_1, X_2 \sim i.i.d. P(\lambda)$. Se considera el estimador de λ , $S_a = a\bar{X}$

- a.- Encuentre el sesgo, varianza y E.C.M. de $S_a(X)$ en función de a y de λ .
- b.- Para λ fijo, encuentre el valor de a que minimize el E.C.M.
- c.- Para a fijo, determine para qué valores de λ , $S_a(X)$ es mejor.

7. Se tiene una muestra aleatoria simple Y_1, \dots, Y_n de una variable aleatoria Y tal que $E(Y) = \mu$, $V(Y) = \sigma^2$. Considere el estimador de μ

$$\hat{\mu} = \sum a_i Y_i .$$

- a) Determinar una condición para que $\hat{\mu}$ sea un estimador insesgado.
- b) Encuentre la varianza de $\hat{\mu}$ y determine los a_i de modo que la varianza sea mínima.
- c) Encuentre los valores de a_i de manera que $ECM(\hat{\mu})$ sea mínimo. Comente sus resultados.

8. La variable aleatoria X está distribuida uniformemente en el intervalo $(0, \theta)$. Se toma una muestra aleatoria de tamaño 1 para estimar θ . Estudie y compare las propiedades de los siguientes estimadores:

$$\hat{\theta}_1 = X_1, \quad \hat{\theta}_2 = 2X_1 .$$

9. Sea $X \sim \mathcal{P}(\lambda)$. Verifique que el estimador de $p_k = P(X = k)$

$$\hat{p}_k = \begin{cases} \frac{r!(n-1)^{r-k}}{k!(r-k)!n^r} & , \quad r = \sum_{i=1}^n X_i \geq k \\ 0 & , \quad r < k \end{cases}$$

es insesgado.

10. Sea Y_1, \dots, Y_n una muestra de la distribución Normal bivariada con media μ y matriz de covarianzas Σ . Escriba la función verosimilitud y encuentre estadígrafos suficientes para μ y Σ .

11. Sea Y_1, \dots, Y_n una muestra de la distribución $N(\mu, \sigma^2)$.

- a) Grafique la función de verosimilitud y su logaritmo si $\sigma^2 = 1$, $\bar{Y} = 10$.
- b) Grafique el logaritmo de la función verosimilitud si $\bar{Y} = 10$, $\sum (Y_i - \bar{Y})^2 = 20$.
- c) Repita (a) si el tamaño de la muestra es 100. Compare los resultados.

12. Sean $X_1, \dots, X_n \sim iid N(\mu, \sigma^2)$ y considere la familia de estimadores

$$\hat{\sigma}_\alpha^2 = \alpha \sum_{i=1}^n (X_i - \bar{X})^2 .$$

Encuentre el valor de α que minimize el E.C.M.

13. Sea Y_1, \dots, Y_n una muestra aleatoria de la variable $X \sim \mathcal{P}(\mu)$ donde μ es un parámetro desconocido. Demuestre que

$$\hat{\mu}_1 = \frac{1}{n} \sum X_i \quad \text{y} \quad \hat{\mu}_2 = \frac{2}{n(n+1)} \sum iX_i$$

son estimadores insesgados de μ . Calcule las varianzas de estos estimadores y compare sus propiedades.

14. Sea Y_1, \dots, Y_n una muestra aleatoria de la variable $Y \sim B(1, \pi)$, donde π es un parámetro desconocido. Considere los estimadores

$$\hat{\pi}_1 = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{y} \quad \hat{\pi}_2 = \frac{1}{2}(X_1 + X_n).$$

- Demuestre que $\hat{\pi}_1$ y $\hat{\pi}_2$ son estimadores insesgados de π .
 - Calcule la varianza de cada estimador y la eficiencia relativa de $\hat{\pi}_1$ respecto de $\hat{\pi}_2$.
 - Encuentre el máximo de la varianza del estimador $\hat{\pi}_1$ para n fijo.
 - Obtenga un estimador insesgado de $V(Y)$ en términos de $\hat{\pi}_1$.
 - Proponga estimadores insesgados para las varianzas de $\hat{\pi}_1$ y $\hat{\pi}_2$.
15. Sea Y_1, \dots, Y_n una muestra aleatoria de la variable Y tal que $E(Y) = \mu$ y $V(Y) = \sigma^2$. Demuestre que
- $S_1^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2$ no es un estimador insesgado de σ^2 .
 - $S_2^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2$ es un estimador insesgado de σ^2 .
 - S_1^2 es un estimador asintóticamente insesgado de σ^2 .
- ¿ Se puede concluir que S_2^2 es mejor estimador que S_1^2 ?.

16. Se tiene una muestra aleatoria de una variable indicatriz que toma valores 1 y 0 con probabilidades π y $(1 - \pi)$ respectivamente. Para estimar π se proponen los estimadores

$$\hat{\pi}_1 = \frac{1}{n} \sum Y_i, \quad \hat{\pi}_2 = \left(\frac{n}{n+2} \right) \hat{\pi}_1 + \frac{1}{n+2}.$$

Si $n = 10$, estudie el comportamiento de la eficiencia relativa de $\hat{\pi}_1$ respecto de $\hat{\pi}_2$ para distintos valores de π .

17. Sea Y_1, \dots, Y_n una muestra aleatoria de la variable Y que tiene distribución Normal con media μ y varianza σ^2 desconocidas. Sean $\bar{Y} = \frac{1}{n} \sum Y_i$ y $S^2 = \frac{1}{n} \sum (Y_i - \bar{Y})^2$ la media y la varianza muestral respectivamente.

- ¿Son estos estimadores consistentes de μ y σ^2 ?.
- Proponga un estimador para el coeficiente de variación $CV = \sigma/\mu$. Mencione sus propiedades.

c) Suponga que $\mu = \sigma^2 = \theta$. Estudie las propiedades del estimador $\hat{\theta} = \bar{Y}$.

18. La variable aleatoria X está distribuida uniformemente en el intervalo $(0, \theta)$. Se toma una muestra aleatoria de tamaño 1 para estimar θ . Estudie y compare las propiedades de los siguientes estimadores:

$$\hat{\theta}_1 = X_1, \quad \hat{\theta}_2 = 2X_1.$$

19. Sean $Y_1, \dots, Y_n \sim U[0, \theta]$. Sea $T = \max(Y_1, \dots, Y_n)$ y considere los estimadores de θ de la forma cT , $c > 0$.

- a) ¿Para qué valor de c , cT es insesgado ?
b) ¿Para qué valor de c , $ECM(cT)$ es mínimo ?.

20. Sean Y_1, \dots, Y_n una muestra aleatoria de una distribución Log-Normal (α, β) . Sea T_n la media geométrica de la muestra:

$$T_n = \exp \left(\frac{1}{n} \sum_{i=1}^n \log Y_i \right)$$

y sea γ la mediana de la distribución de los Y_i .

- a) Demuestre que para muestras grandes T_n converge a γ .
b) Calcule el sesgo de T_n cuando se usa como estimador de γ .

21. Sea $Y_1, \dots, Y_n \sim N(\mu, \sigma^2)$.

- a) Considere el estimador

$$\hat{\sigma}^2 = \frac{\sum (Y_i - \bar{Y})^2}{k}.$$

y encuentre k de manera que $ECM(\hat{\sigma}^2)$ sea mínimo.

- b) Si μ es conocido, demuestre que

$$\hat{\sigma}^2 = \frac{\sum (Y_i - \mu)^2}{n}$$

alcanza la cota de Cramer–Rao. Comente la significancia de este resultado.

22. Encuentre estadígrafos suficientes para estimar θ en base a una muestra de las siguientes distribuciones con a conocido :

- a) $p(x, \theta) = \theta x^{\theta-1}$, $0 < x < 1$, $\theta > 0$.
b) $p(x, \theta) = \theta a x^{a-1} \exp\{-\theta x^a\}$, $x > 0$, $\theta > 0$, $a > 0$.
c) $p(x, \theta) = \frac{\theta a^\theta}{x^{\theta+1}}$, $x > a$, $\theta > 0$, $a > 0$,

23. Tres muestras independientes de tamaños 10, 20 y 20 de una distribución $N(\mu, \sigma^2)$ entregaron las siguientes estimaciones (en base al estimador insesgado usual con μ desconocido) : $s_1^2 = 12,8$, $s_2^2 = 19,6$, $s_3^2 = 20,8$. Proponga un estimador de σ^2 que supere a los tres estimadores propuestos. Demuestre que su estimador es mejor.
24. Sea Y_1, \dots, Y_n una muestra de la distribución Normal bivariada con media μ y matriz de covarianzas Σ . Escriba la función verosimilitud y encuentre estadígrafos suficientes para μ y Σ .
25. Sea Y_1, \dots, Y_n una muestra de la distribución $N(\mu, \sigma^2)$.
 - a) Grafique la función de verosimilitud y su logaritmo si $\sigma^2 = 1$, $\bar{Y} = 10$.
 - b) Grafique el logaritmo de la función verosimilitud si $\bar{Y} = 10$, $\sum (Y_i - \bar{Y})^2 = 20$.
 - c) Repita (a) si el tamaño de la muestra es 100. Compare los resultados.

Métodos de Estimación

1. Sea Y_1, \dots, Y_n una muestra aleatoria de la distribución $Poisson(\mu)$, donde μ es desconocido. Encuentre el estimador máximo verosímil de μ y el estimador máximo verosímil de la probabilidad de observar el valor 0. Calcule las varianzas de estos estimadores y estimaciones máximo verosímiles para estas varianzas.
2. Sea Y_1, \dots, Y_n una muestra aleatoria de la distribución $Binomial(1, \pi)$, donde π es un parámetro desconocido. Calcule los estimadores máximo verosímil y de mínimos cuadrados de π . Calcule un estimador insesgado para las varianzas de estos estimadores.
3. Suponga una muestra de tamaño n de la distribución $Normal(\mu, \sigma^2)$
 - a) Encuentre el estimador máximo verosímil para μ , con σ^2 conocido y el estimador máximo verosímil para σ^2 , con μ conocido.
 - b) Encuentre el estimador máximo verosímil para $\theta = (\mu, \sigma^2)$, ambos desconocidos. Calcule las varianzas de estos estimadores para muestras pequeñas y muestras grandes.
4. Extienda sus resultados para estimar los parámetros de una distribución $Normal$ bivariada a partir de una muestra de tamaño n . En particular obtenga el estimador máximo verosímil para el coeficiente de correlación.
5. Suponga que se eligen al azar n piezas cilíndricas entre las producidas por cierta máquina y se miden sus diámetros y longitudes. Se observó que en n_1 ambas medidas son inferiores a los límites de tolerancia; en n_2 las longitudes son satisfactorias pero no los diámetros; en n_3 los diámetros pero no las longitudes; y en n_4 ninguna de las medidas son satisfactorias (se verifica $n_1 + n_2 + n_3 + n_4 = n$). Cada pieza puede considerarse procedente de una población con distribución $Multinomial(n, \pi_1, \pi_2, \pi_3, \pi_4)$. Donde $\sum_i \pi_i = 1$. ¿Cuáles son las estimaciones máximo verosímiles de los parámetros de esta distribución si $n_1 = 90$, $n_2 = 6$, $n_3 = 3$ y $n_4 = 1$? Aplique el método de χ^2 mínimo y compare los resultados.

6. De cada una de cuatro poblaciones con distribuciones de probabilidad $Normal(\mu_1, \sigma^2), \dots, Normal(\mu_4, \sigma^2)$ respectivamente, se saca una muestra aleatoria de tamaño n . Si

$$\mu_1 = a + b + c, \quad \mu_3 = a - b + c$$

$$\mu_2 = a + b - c, \quad \mu_4 = a - b - c$$

¿Cuáles son los estimadores de máxima verosimilitud de a, b, c y σ^2 ? ¿Cuáles son los estimadores de mínimos cuadrados para estos parámetros?

7. Sea Y_1, \dots, Y_n una muestra aleatoria de una distribución que tiene densidad:

$$f(y, \theta_1, \theta_2) = \frac{1}{\theta_2} e^{-\frac{(y-\theta_1)}{\theta_2}}$$

$$\theta_1 \leq y < \infty, \quad -\infty < \theta_1 < \infty, \quad 0 < \theta_2 < \infty.$$

Encuentre los estimadores máximo verosímil de θ_1 y θ_2 .

8. La duración de una cierta marca de ampolletas tiene una distribución Exponencial con parámetro λ y media $1/\lambda$. Una muestra de 50 ampolletas es ensayada, registrándose cada vez la duración medida en horas :

1000	1010	1125	1181	1230	1250	1265	1270	1275	1285
1327	1340	1348	1350	1365	1370	1380	1400	1406	1409
1430	1445	1448	1502	1505	1507	1510	1515	1521	1525
1540	1560	1580	1595	1602	1605	1608	1610	1620	1630
1710	1720	1730	1780	1810	1815	1860	1940	1984	2000

- a) Grafique la función verosimilitud usando sólo las dos primeras columnas de observaciones y luego todas las observaciones.
- b) Obtenga una estimación de máxima verosimilitud para λ usando la información dada. Estime la varianza del estimador.

9. Suponga que se dispone de una muestra aleatoria de tamaño n de una variable con distribución $Gamma(\alpha, r)$ con función densidad,

$$f(y) = \begin{cases} \frac{\alpha}{\Gamma(r)} (\alpha y)^{r-1} e^{-\alpha y}, & y > 0; \\ 0, & \text{e.o.c.} \end{cases}$$

donde α y r son parámetros desconocidos. Encuentre estimadores para ambos parámetros a partir del método de momentos y el método de máxima verosimilitud.

10. Muestre que el estimador máximo verosímil de θ en la distribución $U[\theta, \theta + 1]$, construido a partir de una muestra aleatoria de tamaño n , no es único.

11. Suponga que la vida útil de un cierto tipo de ampolletas tiene distribución exponencial de parámetro λ , con λ desconocido. Suponga que n ampolletas son probadas durante T horas y que x_0 de ellas falla antes de completarse el período de T horas. Encuentre el estimador máximo verosímil de λ basado en los datos.
12. Suponga que en un proceso productivo se desea estimar la tasa de falla π .
- Si se inspeccionan n elementos de este proceso y se encuentran r defectuosos, ¿cuál es el estimador máximo verosímil de π ? ¿es insesgado?.
 - Si se decide, ya avanzado el experimento, continuar realizándolo hasta obtener el r -ésimo defectuoso, cosa que ocurre en el n -ésimo elemento ¿cuál es el estimador máximo verosímil de π ? ¿es insesgado?.
13. Sea $f_\theta(x) = \theta x^{\theta-1}$, $0 < x < 1$ y

$$\Pi(\theta) = \frac{\alpha(\alpha\theta)^{\beta-1}e^{-\alpha\theta}}{\Gamma(\beta)} \quad , \quad \theta > 0 .$$

Encuentre $E(\theta|X)$.