SOLUCION EJERCICIOS AUXILIAR NUMERO DOS MA34B

Auxiliar: Diego Díaz Espinoza, Profesor: Alexis Peña.

August 15, 2007

1. Dada la muestra aleatoria: $\bar{X} = (x_1, ..., x_n)$ tal que $x_i \sim N(\mu_1, \sigma_1^2)$ encontrar el EMM para μ_1 y σ_1^2. Dado que se quieren calcular dos parámetros, se necesita encontrar $\mu_1(\hat{\mu}) = E(x)$ y $\mu_2(\hat{\mu}) = E(x^2)$ e igualar a $m_1 = \frac{\sum_{i=1}^n x_i}{n}$ y $m_2 = \frac{\sum_{i=1}^n x_i^2}{n}$ respectivamente para encontrar los estimadores de momentos.
Primer método: Usando integración. Se usa integración para calcular las esperanzas. Así:

$$E(x) = \int_{-\infty}^{+\infty} x \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{1}{2\sigma^2}(x-\mu)^2} dx$$

Haciendo el cambio de variable $(x - \mu) = r \Rightarrow dx = dr$ se tiene

$$E(x) = \int_{-\infty}^{+\infty} (r+\mu) \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-1}{2\sigma^2}r^2} dr = \frac{1}{\sqrt{2\pi}\sigma} \left[\int_{-\infty}^{+\infty} r e^{\frac{-1}{2\sigma^2}r^2} dr + \int_{-\infty}^{+\infty} \mu e^{\frac{-1}{2\sigma^2}r^2} dr \right]$$

$$= \frac{1}{\sqrt{2\pi}\sigma} \left[-2\sigma^2 e^{\frac{-1}{2\sigma^2}r^2} \mid_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} \mu e^{\frac{-1}{2\sigma^2}r^2} dr \right] = \frac{1}{\sqrt{2\pi}\sigma} \left[0 + \int_{-\infty}^{+\infty} \mu e^{\frac{-1}{2\sigma^2}r^2} dr^{(*)} \right] = \frac{1}{\sqrt{2\pi}\sigma} \left[\frac{2\mu\sqrt{\pi}}{2\frac{1}{\sqrt{2}\sigma}} \right] = \mu$$

Ahora veamos $E(x^2)$:

$$E(x^2) = \int_{-\infty}^{+\infty} x^2 \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{1}{2\sigma^2}(x-\mu)^2} dx$$

Usando el mismo cambio de variable se tiene:

$$E(x^2) = \int_{-\infty}^{+\infty} (r+\mu)^2 \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-1}{2\sigma^2}r^2} dr = \frac{1}{\sqrt{2\pi}\sigma} \left[\int_{-\infty}^{+\infty} (r^2 + 2r\mu + \mu^2) e^{\frac{-1}{2\sigma^2}r^2} dr \right]$$

$$= \frac{1}{\sqrt{2\pi}\sigma}\left[\int_{-\infty}^{+\infty} r^2 e^{\frac{-1}{2\sigma^2}r^2}dr + 2\mu\int_{-\infty}^{+\infty} re^{\frac{-1}{2\sigma^2}r^2}dr + \mu^2\int_{-\infty}^{+\infty} e^{\frac{-1}{2\sigma^2}r^2}dr\right]$$

$$= \frac{1}{\sqrt{2\pi}\sigma}\left[\int_{-\infty}^{+\infty} r^2 e^{\frac{-1}{2\sigma^2}r^2}dr + 0 + \int_{-\infty}^{+\infty} \mu^2 e^{\frac{-1}{2\sigma^2}r^2}dr\right] = \frac{1}{\sqrt{2\pi}\sigma}\left[\int_{-\infty}^{+\infty} r^2 e^{\frac{-1}{2\sigma^2}r^2}dr + 0 + \mu^2\right]$$

$$= \frac{1}{\sqrt{2\pi}\sigma}\left[\frac{\sigma^2\sqrt{\pi}}{\frac{1}{\sqrt{2}\sigma}}^{(*)} + 0 + \frac{\mu^2\sqrt{\pi}}{\frac{1}{\sqrt{2}\sigma}}\right]$$

En donde las integrales marcadas con (*) deben verse en tablas pues no tienen primitiva conocida (aunque pueden resolverse usando integración compleja o pasando a funciones de taylor...quizá).

Segundo método: usando propiedades: $E(x) = \mu$(por tabla) y usando las propiedades de la varianza se tiene que

$$Var(x) = E(x^2) - E(x)^2 \Rightarrow E(x^2) = Var(x) + E(x)^2 = \sigma^2 + \mu^2$$

Luego, teniendo las esperanzas se igualan los momentos poblacionales ($\mu_1(\hat{\mu})$ y $\mu_2(\hat{\mu})$) con los momentos muestrales (m_1y m_2) (notar que los momentos muestrales dependen sólo de la muestra aleatoria y los poblacionales involucran a la población completa, luego lo que se está haciendo es imponer que la muestra es representativa de la población), con esto se obtienen las ecuaciones EMM:

$$\mu_1(\hat{\mu}) = m_1 \Rightarrow \hat{\mu_1} = \frac{\sum_{i=1}^n x_i}{n} = \bar{X}$$

$$\mu_1(\hat{\mu}) = m_1 \Rightarrow \hat{\sigma_1^2} = \frac{\sum_{i=1}^n x_i^2}{n} + \hat{\mu} = \frac{\sum_{i=1}^n x_i^2}{n} + \bar{X} = \frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n}$$

Así los parámetros estimados son:

$$\hat{\mu_1} = \bar{X}$$

$$\hat{\sigma_1^2} = \frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n}$$

Consideremos ahora otra muestra que sigue la misma distribución normal pero distintos parámetros, esto es: $y_i \sim N(\mu_2, \sigma_2^2)$ y veamos cómo se distribuye una muestra aleatoria $Z = (z_1, ..., z_n) = (x_1 + y_1, ..., x_n + y_n)$. Primero debemos buscar cómo se distribuye z:

Primer método para calcular z, usando la f.g.m.(función generadora de momentos):

$$E(e^{zt}) = E(e^{(x+y)t}) = E(e^{xt}e^{yt})^{(*)} = E(e^{xt})E(e^{yt}) = fgm(x)fgm(y)^{(**)} = e^{(t\mu_1+\frac{\sigma_1^2t^2}{2})}e^{(t\mu_2+\frac{\sigma_2^2t^2}{2})}$$

$$e^{(t\mu_1+\frac{\sigma_1^2t^2}{2})+(t\mu_2+\frac{\sigma_2^2t^2}{2})} = e^{(t\mu_1+t\mu_2+\frac{\sigma_1^2t^2}{2}+\frac{\sigma_2^2t^2}{2})} = e^{[t(\mu_1+t\mu_2)+t^2(\frac{\sigma_1^2+\sigma_2^2}{2})]}$$

Y como hay unicidad entre fgm y distribución, se puede concluir que $z \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$./
Otra forma más astuta de llegar al mismo resultado es:

$$E(z_i) = E(x_i + y_i) = E(x_i) + E(y_i) =^{(*)} \mu_1 + \mu_2$$

$$Var(z_i) = Var(x_i + y_i) =^{(*)} Var(x_i) + Var(y_i) = \sigma_1^2 + \sigma_2^2$$

Donde (*) se debe a que $x_i \perp y_i$ y (**) se extrae de las tablas de distribuciones o calculando

$$E(e^{xt}) = \frac{1}{\sqrt{2\pi}\sigma}\left[\int_{-\infty}^{+\infty} e^{xt}e^{\frac{1}{2\sigma^2}(x-\mu)^2}dx\right] = \frac{1}{\sqrt{2\pi}\sigma}\left[\int_{-\infty}^{+\infty} e^{xt+\frac{1}{2\sigma^2}(x-\mu)^2}dx\right]$$

Que es relativamente fácil de integrar (ejercicio).
Ahora que tenemos la distribución de Z sólo nos queda estimar sus parámetros, el caso es el mismo que para una normal y en realidad el ejemplo se hizo sólo para introducir el uso de fgm y propiedades.

2. ESTUDIO DE CASO UNO.
Primero calcular las frecuencias del número de camiones (cuántas veces pasan 0 camiones, cuántas veces pasa 1 camión, etc, para cada día). Luego se grafican esos resultados (gráfico 1) y se puede estimar una distribución. En este caso vemos que puede ser Poisson (notar que la gráfica es decreciente y además los datos son enteros, esto es: información de una población discreta). Luego estimamos el parámetro dela Poisson con los datos empíricos del problema y vemos si se ajusta a ellos contrastando los gráficos teóricos y experimentales (gráfico 2). En este caso además se plantea el uso del siguiente estimador para el parámetro de la Poisson: $\lambda = -ln(\frac{X_0}{n})$ donde X_0es la cantidad de ceros de la muestra (cuántos días pasaron 0 camiones por el puente) con esto obtenemos una tercera aproximación para el problema pero con otro estimador (gráfico 3)...Qué hay del estimador Máximo Verosímil?...ejercicio.

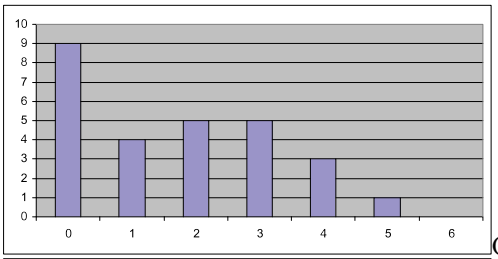


Gráfico 1.

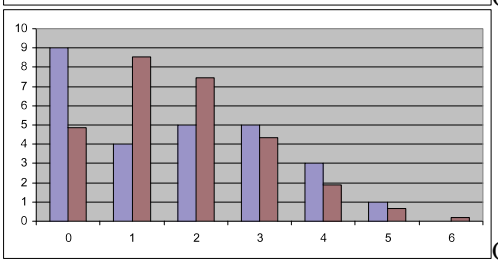


Gráfico 2.

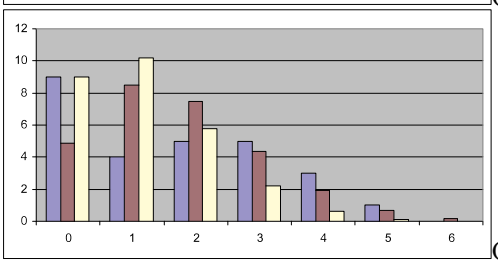


Gráfico 3.

Cálculo del EMM para la distribución de Poisson:
Usando las tablas estadísticas se tiene que el momento poblacional es $\mu_1(\hat{\lambda}) = E(x) = \lambda$ Por otro lado, el momento muestral es $m_1 = \frac{\sum_{i=1}^{n} x_i}{n}$.
Con esto el EMM queda determinado por:

$$\hat{\lambda} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{X}$$

3. Ejercicios resueltos EMV.

Poisson: Debemos estimar λ:
Primero calculemos la función de verosimilitud:

$$L(\lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} = \frac{\lambda^{\sum x_i} e^{-\sum \lambda}}{\prod_{i=1}^{n} x_i!} = \frac{\lambda^{\sum x_i} e^{-n\lambda}}{\prod_{i=1}^{n} x_i!}$$

Maximizando esta función, se tiene:

$$\frac{\partial L(\lambda)}{\partial \lambda} = \frac{\sum x_i \lambda^{\sum x_i - 1} e^{-n\lambda} - n\lambda^{\sum x_i} e^{-n\lambda}}{\prod_{i=1}^{n} x_i!}$$

$$\frac{\partial L(\lambda)}{\partial \lambda} = 0 \Leftrightarrow \frac{\sum x_i \lambda^{\sum x_i - 1} e^{-n\lambda} - n\lambda^{\sum x_i} e^{-n\lambda}}{\prod_{i=1}^{n} x_i!} = 0 \Leftrightarrow e^{-n\lambda}\lambda^{\sum x_i}\left[\sum x_i \lambda^{-1} - n\right] = 0$$

Pero $e^{-n\lambda}\lambda^{\sum x_i} \neq 0 \quad \forall x_i, \lambda > 0$. Luego se tiene el EMV para Poisson:

$$\hat{\lambda} = \frac{\sum x_i}{n} = \bar{X}$$

Luego el EMM y el EMV coinciden. Además el estimador es insesgado.
Los ejercicios para la exponencial y la bernoulli los ve el profesor en clases.
Veamos un ejercicio de control (control uno primavera 2006):Dada la siguiente función de densidad para una muestra aleatoria:

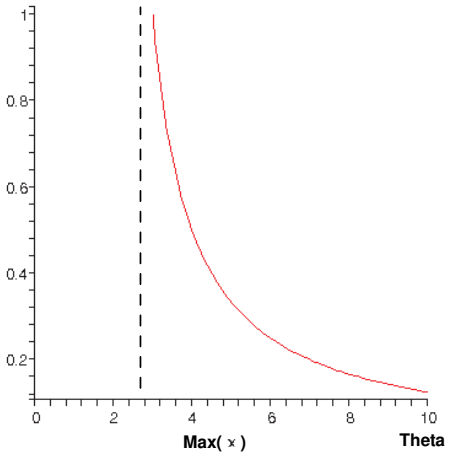
$$f(x|(\theta, \alpha)) = \frac{(\alpha+1)x^{\alpha}}{\theta^{\alpha+1}} \qquad 0 < x < \theta, \quad \alpha conocido$$

Se pide calcular dos cosas:
Primero el EMV:
Para ello calculamos la función de verosimilitud y luego maximizamos:

$$L(\theta) = \frac{(\alpha+1)^n \left(\prod x_i\right)^{\alpha}}{\theta^{n\alpha+n}}$$

$$\frac{\partial L(\theta)}{\partial \theta} = \frac{-(\alpha+1)^n (n\alpha+n)\left(\prod x_i\right)^{\alpha}}{\theta^{n\alpha+n-1}} \neq 0$$

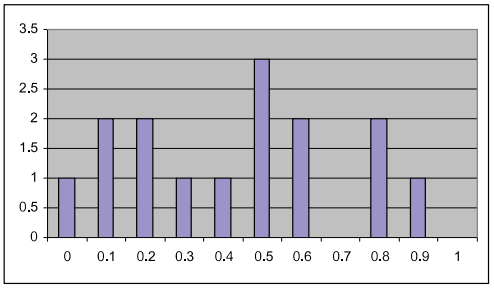
Luego para encontrar el EMV debemos verlo gráficamente:

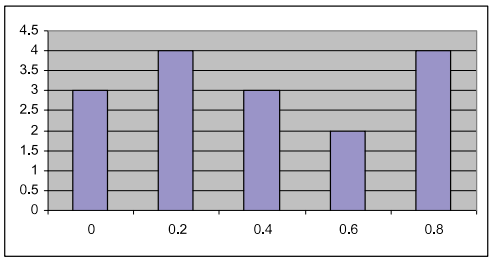


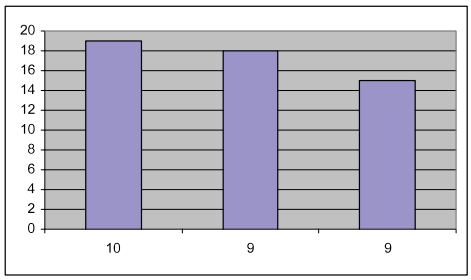
Luego el EMV es $\hat{\theta} = Max\{x_i\}$. Es insesgado? si no lo es construya uno a partir de este que sí lo sea (ejercicio).

4. ESTUDIO DE CASO DOS.

   Usando el mismo método que se usó en el caso uno se tendrán los gráficos correspondientes que indican una forma Uniforme para la distribución (en cada uno de ellos se ha agrupado en distintas categorías de frecuencias y para cada una de ellas no cabe otra que una uniforme). Sabemos que el estimador para una uniforme es $T = Max\{x_i\}$ (ver apuntes de cátedra).







NOTA: Notar que esta forma de estimar distribuciones de datos es muy débil, en la última parte del curso verán un test para contrastar distribuciones sobre los datos empíricos.