Inferencia estadística
Tarea 2

J. Antonio García, jose.ramirez@cimat.mx

19 de octubre de 2017

1. Considera la v.a. X con densidad:

$$f(x) = \begin{cases} 3x^-4, & x \geq 1 \\ 0 & en \quad otro \quad caso \end{cases}$$

Definamos la v.a. $Y = X^{-1}$. Encuentra la función de densidad de Y. Verifica que efectivamente X y Y no son independientes (usa valores esperados).

_*_*_

Dado que la función $Y(x) = X^{-1}$ es inyectiva en $[1, \infty]$ (de hecho es monótona decreciente) podemos emplear el resultado 2.12 del texto que seguimos.
Tenemos que:

$$\begin{aligned} F_y(y) &= P(Y \leq y) \\ &= P(\frac{1}{x} \leq y) \\ &= P(\frac{1}{y} \leq x) \\ &= 1 - P(x \leq \frac{1}{y}) \\ &= 1 - \int_1^{\frac{1}{y}} 3x^{-4} dx \\ &= 1 - 3\left(\frac{x^-3}{-3}\right)\Big|_1^{\frac{1}{y}} \\ &= 1 + x^{-3}\Big|_1^{\frac{1}{y}} \\ &= 1 + \left((\frac{1}{y})^{-3} - 1\right) \\ &= (\frac{1}{y})^{-3} \\ &= y^3 \end{aligned}$$

Con lo cual es muy fácil encontrar $f_y(y)$

$$f_y(y) = F'y(y) = (y^3)' = 3y^2$$

Para probar que $X \not\perp Y$ notemos que $\mu_y \neq \frac{1}{\mu_x}$.
Pues $\mu_y = \int_1^\infty \frac{1}{x} 3x^{-4} = \int_1^\infty 3x^{-5} = -\frac{3}{4}x^{-4}|_1^\infty = 4/3$.
Mientras que $\mu_x = \int_1^\infty x3x^{-4} = 3\int_1^\infty x^{-3} = \frac{-3}{2}x^{-2}|_1^\infty = \frac{3}{2}$

2. Sean $\bar{X}_n$ y S_n^2 la media y la varianza, respectivamente, de $X_1, \ldots, X_n$. Supón que te llega otra observación. Muestra que

a) $\bar{X}_{n+1} = \frac{X_{n+1}+n\bar{X}_n}{n+1}$

Recordemos que $\bar{X}_n = \sum_{i=1}^{n} X_i/n$ y $S_n^2 = \sum_{i=1}^{n} \frac{(X_i-\bar{X}_n)^2}{n-1}$. Entonces:

$\bar{X}_{n+1} = (\frac{1}{n}) \sum_{i=1}^{n+1} X_i = (\frac{1}{n+1}(\sum_{i=1} nX_i + X_{n+1}) = (\frac{1}{n+1})(n\bar{X}_n + X_{n+1}) = \frac{x_{n+1}+n\bar{X}_n}{n+1}$

b) $nS_{n+1}^2 = (n-1])S_n^2 + (\frac{n}{n+1})(X_{n+1} - \bar{X}_n)^2$

$$nS_{n+1}^2 = n(\frac{\sum_{i=1}^{n+1}(X_i - \bar{X}_{n+1})^2}{n}) = \sum_{i=1}^{n}(x_i - \bar{X}_{n+1})^2 + (x_{n+1} - \bar{X}_{n+1})^2$$

$$= \sum_{i=1}^{n}(x_i - \frac{1}{n+1}(X_{n+1} + n\bar{X}_n))^2 + (X_{n+1} - \frac{1}{n+1}(X_{n+1} + n\bar{X}_n))^2$$

$$= \sum_{i=1}^{n}\left(x_i + (-\bar{X}_n + \bar{X}_n) - \frac{1}{n+1}(X_{n+1} + n\bar{X}_n)\right)^2 + (\frac{1}{n+1}((n+1)X_{n+1} - X_{n+1} - n\bar{X}_n))^2$$

$$= \sum_{i=1}^{n}\left((X_i - \bar{X}_n) + \bar{X}_n - \frac{1}{n+1}(X_{n+1} + n\bar{X}_n)\right)^2 + (\frac{nX_{n+1} - n\bar{X}_n}{n+1})^2$$

$$= \sum_{i=1}^{n}\left((X_i - \bar{X}_n) + \frac{(n+1)\bar{X}_n - X_{n+1} - n\bar{X}_n}{n+1}\right)^2 + \frac{n^2}{(n+1)^2}(X_{n+1} - \bar{X}_n)^2$$

$$= \sum_{i=1}^{n}\left((X_i - \bar{X}_n) + \frac{\bar{X}_n - X_{n+1}}{n+1}\right)^2 + \frac{n^2}{(n+1)^2}(X_{n+1} - \bar{X}_n)^2$$

$$= \sum_{i=1}^{n}\left((X_i - \bar{X}_n)^2 + 2(X_i - \bar{X}_n)\frac{\bar{X}_n - X_{n+1}}{n+1} + \left(\frac{\bar{X}_n - X_{n+1}}{n+1}\right)^2\right) + \frac{n^2}{(n+1)^2}(X_{n+1} - \bar{X}_n)^2$$

$$= \sum_{i=1}^{n}(X_i - \bar{X}_n)^2 + 2\sum_{i=1}^{n}(X_i - \bar{X}_n)\frac{\bar{X}_n - X_{n+1}}{n+1} + \sum_{i=1}^{n}\left(\frac{\bar{X}_n - X_{n+1}}{n+1}\right)^2 + \frac{n^2}{(n+1)^2}(X_{n+1} - \bar{X}_n)^2$$

$$= (n-1)S_n^2 + 2(\sum_{i=1}^{n} X_i - \sum_{i=1}^{n}\bar{X}_n)\frac{\bar{X}_n - X_{n+1}}{n+1} + \frac{n}{(n+1)^2}(\bar{X}_n - X_{n+1})^2 + \frac{n^2}{(n+1)^2}(X_{n+1} - \bar{X}_n)^2$$

$$= (n-1)S_n^2 + 2(n\bar{X}_n - n\bar{X}_n)\frac{\bar{X}_n - X_{n+1}}{n+1} + \frac{n}{(n+1)^2}(\bar{X}_n - X_{n+1})^2 + \frac{n^2}{(n+1)^2}(X_{n+1} - \bar{X}_n)^2$$

$$= (n-1)S_n^2 + \frac{n}{(n+1)^2}((-1)^2X_{n+1} - \bar{X}_n)^2 + \frac{n^2}{(n+1)^2}(X_{n+1} - \bar{X}_n)^2$$

$$= (n-1)S_n^2 + \frac{n^2+n}{(n+1)^2}(X_{n+1} - \bar{X}_n)^2$$

$$= (n-1)S_n^2 + \frac{n}{n+1}(X_{n+1} - \bar{X}_n)^2$$

3. Pedro y Ana quieren recortar un rectángulo de papel. Como ambos estudiaron probabilidad, calculan la forma exacta del rectángulo generando v.a. positivas U de la siguiente forma. Pedro es flojo, y genera una sola v.a. U, y luego corta un cuadrado con ésta longitud en cada lado. A Ana le gusta la diversidad, así que genera dos v.a. independientes U, y recorta un rectángulo con ancho igual a la primera v.a. y largo igual a la segunda v.a.

   a) ¿Las áreas de los rectángulos de Pedro y Ana serán diferentes en promedio ?
   _*_*_*_
   Sí lo son

   b) ¿Cuál rectángulo, el de Pedro o Ana se esperaría que fuera más grande ?
   _*_*_*_
   Consideremos que las variables generadas se comportan como $U(0, b)$ Con el teorema de Pitágoras podemos deducir que, en el esquema de Pedro, U es la v.a. que representa la diagonal de sus rectángulos entonces la función $P(U) = \frac{u^2}{2}$ indica el área de los rectángulos de Pedro.
   Por su parte el área de los rectángulos de Ana está dada por la función $A(U_1, U_2) = U_1U_2$ donde $U_i \sim U(0, b)$ y $U_1 \perp U_2$.
   Calculamos el valor esperado para ambas funciones de v.a's para mostrar que los rectángulos de

Pedro serán más pequeños.

$$\mu_P = \int_0^b P(u)f(u) = \int_0^b \frac{u^2}{2}(\frac{1}{b}) = \frac{1}{2b}\frac{u^3}{3}|_0^b = \frac{1}{6b}(b^3) = \frac{b^2}{6}$$

Por su parte los de Ana

$$\mu_A = \int_0^b\int_0^b A(u_1,u_2)f(u_1,u_2) = \int_0^b\int_0^b u_1u_2f(u_1,u_2) = \int_0^b\int_0^b u_1u_2f(u_1)f(u_2) = \int_0^b u_1f(u_1)\int_0^b u_2f(u_2)$$
$$= \mu_{u_1}\mu_{u_2} = \frac{b}{2}\frac{b}{2} = \frac{b^2}{4}$$

4. Considera T una v.a. exponencial, que como ya vimos, tiene una función de densidad $f(t) = \lambda e^{-\lambda t}$ y $E(T) = 1/\lambda$. Esta v.a. tiene dos propiedades importantes, que ya se mencionaron en clase:

   - El valor mínimo de variables aleatorias exponenciales es exponencial cuyo parámetro es la suma de los parámetros individuales.
   - Para dos v.a. exponenciales T, Q, con $T > Q$, la resta $T - Q$ es exponencial con parámetro λ (pérdida de memoria).

   Considera una muestra de tamaño n de v.a. exponenciales X. Sea $X_{(1)}$ el valor más pequeño de la muestra y $X_{(n)}$ el más grande. Define el Rango $R_n = X_{(n)} - X_{(1)}$, que es una v.a. positiva

   a) Usa las propiedades descritas para calcular $E(R_n)$.

   _*_*_

   Tenemos que si fijamos el tiempo en que ocurre el mínimo entonces el tiempo entre este momento y $X_{(n)}$, es decir $X_{(2)} - X_{(1)}$, equivale a esperar el mínimo de las restantes (utilizando la propiedad de perdida de memoria)$n-1$ v.a., X con distribución $Exp(\lambda)$, análogamente el tiempo entre $X_{(3)}$ y $X_{(2)}$ después de fijar $X_{(2)}$ corresponde al tiempo de observar el mínimo en la muestra de las otras $n-2$ v.a. así hasta que ocurra $X_{(n)}$ con lo que el tiempo de espera entre $X_{(n)}$ y $X_{(n-1)}$ equivale a esperar el mínimo de una sola v.a. por lo que si llamamos $Y_n = X_{(n)} - X_{(n-1)}$ al tiempo de espera entre los estadísticos de orden consecutivos, y cuidamos que $Y_1 = Y_{(1)}$, entonces podemos ver que el tiempo de espera hasta que observamos el valor $X_{(n)}$ es el mismo que $\sum_{i=1}^n Y_n$, además $Y_i \perp Y_j, i \neq j$ Como vimos en clase el mínimo de una muestra de tamaño k que proviene de una v.a. con distribución $Exp(\lambda)$ se distribuye $Exp(k\lambda)$ por lo que $Y_k|Y_1 \sim Exp((n-i)\lambda)$.

   Entonces
   $$\mathcal{E}(X_{(n)} - X_{(1)}) = \mathcal{E}(\sum_{i=1}^n Y_i) = \sum_{i=1}^n \mathcal{E}(Y_i) = \sum_{i=1}^n (\frac{1}{\lambda(n-i)}) = \frac{1}{\lambda}\sum_{j=1}^n (\frac{1}{(j)})$$

   b) Encuentra la distribución de R_n (puedes obtenerla sin hacer cálculos explícitos).

   _*_*_

   Siguiendo la idea que desarrollamos en el inciso anterior tenemos que:

   $$\begin{aligned} P(X_{(n)} - X_{(1)} \leq x) &= P\left((Y_n - Y_1) \leq x|Y_1\right) = P\left(Y_n \leq x + Y_1|Y_1\right)\right) \\ &= P((Y_n \leq x) = P(X_{(n)} \leq x) \\ &= 1 - (1 - F_x(x))^{n-1} \end{aligned}$$

   Donde F_x es la acumulada de una v.a. con distribución $Exp(\lambda)$, por lo que $R_n \sim min(x_1, x_2, \ldots, x_{n-1}$ es decir $R_n \sim Exp((n-1)\lambda)$

5. Lee el capítulo 6 del libro de Wasserman, Convergence of random Variables.

6. Sean $X_, \ldots$ una secuencia de v.a. iid $Uniform(-1,1)$. Define $T_n = \frac{1}{n}\sum_{i=1}^n X_i^2$. Queremos mostrar que, para algun a y cualquier $\epsilon > 0$

$$\lim_{n\to\infty} P(|T_n - a| \geq \epsilon) = 0 \tag{1}$$

Explica cómo puede ser esto cierto. Determina el valor de a.

_*_*_

La ecuación 1 quiere decir que el valor de la variable aleatoria T_n converge a un valor real, esto quiere decir que intuitivamente el valor de T_n se concentra alrededor de un punto (cuando n es grande) o bien que la distribución de T_n se concentra alrededor del valor a.

Para determinar el valor de a podemos usar la ley débil de los grandes números y proponer $a = \mu_{T_n}$, para lo cual se requiere obtener primero f_{X^2} :

$$\begin{aligned} F_{X^2}(y) &= P(Y \leq y) = P(x^2 \leq y) = P(x \leq y^{1/2}) \\ &= \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{1+1} dx = \frac{1}{2} 2\sqrt{y} = \sqrt{y} \end{aligned}$$

Derivando

$$f_{X^2}(y) = (\sqrt{y})' = \frac{1}{2} y^{-1/2}$$

Y calculamos el valor de $a = \mu_{T_n}$

$$\begin{aligned} \mu_{T_n} &= \mathbf{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right) = \frac{1}{n}\mathbf{E}\left(\sum_{i=1}^{n} X_i^2\right) = \frac{1}{n}\left(\sum_{i=1}^{n} \mathbf{E}(X_i^2)\right) = \frac{n}{n}\mathbf{E}(X_i^2) = \mu_{X^2} \\ &= \frac{1}{2}\int_0^1 yy^{-1/2} = \frac{1}{2} y^{1/2} = \frac{1}{2}\frac{y^{3/2}}{3/2} = \frac{1}{3} y^{3/2}|_0^1 = \frac{1}{3} \end{aligned}$$

7. Un contador quiere simplificar su contabilidad redondeando al entero más próximo, por ejemplo 99.53 y 100.46 ambos a 100. ¿Cuál es el efecto acumulativo de hacer esto si hay, digamos, 100 cantidades? Para analizar esto, modelamos los errores de redondeo mediante 100 v.a. $Uniform(-0.5, 0.5)$ independientes $X_1, \ldots, X_{100}$

   a) Calcula el promedio y varianza de X.

   _*_*_

   Definamos la v.a. $X = \sum_{i=1}^{10} X_i$, tenemos que:
   $\mu_X = \mathbf{E}(X) = \mathbf{E}\left(\sum_{i=1}^{100} X_i\right) = \sum_{i=1}^{100} (\mathbf{E}(X_i)) = 100\_X_i = 100(0) = 0$
   Y también
   $\sigma_X^2 = Var(\sum_{i=1}^{100} X_i) = \sum_{i=1}^{100} Var(X_i) = 100(\frac{1}{12}) = \frac{100}{12} = \frac{50}{6}$

   b) Usa la desigualdad de Chevyshev para calcular un límite superior para la probabilidad. $P(|X_1 + X_2 + \cdots + X_{100}| > 10)$ de que el error de redondeo acumulado excederá \$10.
   _*_*_

$$\begin{aligned} P(|X_1 + \cdots + X_{100}| > 10) = P(|X| > 10) = P(|x - \mu_X| > 10) = P(|x - \mu_X| \geq 10) \\ \leq \frac{\sigma_X^2}{10^2} = \frac{\frac{50}{6}}{100} = \frac{50}{600} = \frac{1}{12} \approx 0.08\bar{3} \end{aligned}$$

   c) Usa el teorema del límite central para obtener una mejor estimación de esa probabilidad.

   _*_*_

   Ahora definamos la v.a. $Y = \frac{X}{n}$, por lo que $\mu_Y = \frac{1}{n}\mu_X = 0$ y $\sigma_Y = Var(\frac{1}{n}X) = \frac{1}{n^2}\frac{n}{12} = \frac{1}{120}$.
   Si consideramos que 100 es un tamaño grande muestra entonces, por el teorema del límite central:

$$\frac{\sqrt{n}(Y - \mu_Y)}{\sigma} \sim Z$$

   Donde Z $N(0, 1)$ Por lo que el problema es equivalente a:

$$\begin{aligned} P\left(\frac{\sqrt{n}(Y - \mu_Y)}{\sigma} \geq \frac{10(10-0)}{\frac{1}{\sqrt{120}}}\right) = P\left(\frac{\sqrt{n}(Y - \mu_Y)}{\sigma} \geq \frac{100}{\frac{1}{\sqrt{120}}}\right) = P\left(\frac{\sqrt{n}(Y - \mu_Y)}{\sigma} \geq 100\sqrt{120}\right) \\ = 1 - P\left(\frac{\sqrt{n}(Y - \mu_Y)}{\sigma} \leq 100\sqrt{120}\right) \approx 1 - P(Z \leq 1095) \end{aligned}$$

   Como Z se distribuye normal estándar, entonces la igualdad anterior es prácticamente cero, por lo que al emplear el teorema del límite central la estimación de la probabilidad mejora notoriamente.