

# Métodos Estadísticos

## Tarea II

Rubén Pérez Palacios Lic. Computación Matemática  
Profesora: Dra. Eloísa Díaz Francés Murguía

27 de marzo de 2021

### Problemas

1. Considera la distribución normal para  $X$ , la variable aleatoria que representa el tiempo en segundos sin respirar de los alumnos del grupo. Los parametros son  $(\mu, \sigma)$  donde la media es  $E(X) = \mu$  y la varianza es  $V(X) = \sigma^2$ . Para la muestra observada de  $n$  tiempos sin respirar del grupo, denotada por  $x_1, \dots, x_n$  da las expresiones de unos estimadores de momentos para los parametros normales  $(\mu, \sigma)$ . Calcula los valores particulares que toman para la muestra observada de datos. Recuerda usar los primeros momentos que tengan expresiones más simples. Las expresiones de unos estimadores de momentos para los parametros normales  $(\mu, \sigma)$  son

$$\tilde{\mu} = \mathbb{E}[X] = \frac{\sum_{i=1}^n x_i}{n}, \quad \tilde{\sigma} = \sqrt{\text{Var}(X)} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n}}.$$

De los cuales sus valores particulares para la muestra observada de datos es

$$\mu = 1, \quad \sigma = 0,2$$

2. Da la expresión de la densidad normal conjunta para la muestra de  $n$  variables aleatorias independientes e idénticamente distribuidas como una Normal  $(\mu, \sigma)$ , factorizándola como hizo Fisher,

$$f(\vec{x}; \vec{\theta}) = h(\vec{x}) g(T(\vec{x}); \theta),$$

de manera que identifiques a las dos estadísticas suficientes  $t_1$  y  $t_2$  para  $(\mu, \sigma)$ . Da sus expresiones y calcúlalas para la muestra de tiempos de respiración. Los estimadores de momentos que calculaste en el problema anterior, exprésalos en terminos de las estadísticas suficientes  $t_1$  y  $t_2$ .

Recordemos que la función de densidad de una Normal  $(\mu, \sigma^2)$  es (se omite la indicadora puesto que esta es sobre todos los números reales)

$$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Por lo que la densidad conjunta de  $n$  variables aleatorias independientes e idénticamente distribuidas como una Normal  $(\mu, \sigma)$  es

$$f(\vec{x}; \vec{\theta}) = \prod_{i=1}^n f(x_i; \mu, \sigma^2) = \left( \frac{1}{\sigma\sqrt{2\pi}} \right)^n e^{-\frac{\sum_{i=1}^n x_i^2 - 2\mu \sum_{i=1}^n x_i + n\mu^2}{2}},$$

por el Teorema de la Factorización de Fisher  $\vec{t} = (t_1 = \sum_{i=1}^n x_i^2, \sum_{i=1}^n x_i)$  es el vector de estadísticas suficientes para  $(\mu, \sigma)$ , para los cuales sus valores particulares para la muestra de tiempos de respiración son

$$t_1 = n, \quad t_2 = 0,25.$$

Por último los estimadores de momentos en terminos de las estadísticas suficientes son

$$\mu = t_2/n, \quad \sigma = \sqrt{t_1 + \left( \frac{1-2n}{n^2} \right) t_2^2}.$$

3. Considera ahora las siguientes distribuciones y da la expresión de unos estimadores de momentos. Para ello da explícitamente los valores de los tres siguientes momentos teóricos para tu distribución asignada:  $\mathbb{E}[X]$ ,  $Var(X)$ ,  $\mathbb{E}[X^2]$ . Además calcula y reporta en tu tarea los valores que toman tus estimadores de momentos con los datos observados de tiempos sin respirar.

Para todos los ejercicios se hará uso que

$$\mathbb{E}[X^2] = Var(X) + \mathbb{E}[X]^2,$$

esto puesto que la definición de  $Var$  es

$$Var(x) = \mathbb{E}\left((X - \mathbb{E}X)^2\right).$$

- a) Gamma  $(\alpha, \beta)$  donde  $\alpha$  es parámetro de forma y  $\beta$  es de escala. Por ello  $\mathbb{E}[X] = \alpha\beta$ ,  $Var(X) = \alpha\beta^2$  y la densidad es

$$f(x; \alpha, \beta) = \frac{x^{\alpha-1}}{\Gamma(\alpha)\beta^\alpha} \exp\left(-\frac{x}{\beta}\right) \mathbb{1}_{(-\infty, \infty)}(x).$$

Los momentos teóricos son

- $\mathbb{E}[X] = \sigma\beta.$
- $Var(X) = \sigma\beta^2.$
- $\mathbb{E}[X^2] = \beta^2(\sigma + \sigma^2).$

Por lo tanto los estimadores de momentos son

- $\beta = \frac{Var(X)}{\mathbb{E}[X]} = \frac{\sum_{i=1}^n (x_i - \bar{X})^2}{\sum_{i=1}^n x_i}$
- $\sigma = \frac{\mathbb{E}[X]^2}{Var(X)} = \frac{(\sum_{i=1}^n x_i)^2}{n \sum_{i=1}^n (x_i - \bar{X})^2}$

- b) Logística  $(\mu, \sigma)$  donde  $\mu$  es parámetro de localización y  $\sigma$  es de escala. La función de densidad es

$$f(x; \mu, \sigma) = \frac{\exp\left(-\frac{x-\mu}{\sigma}\right)}{\sigma \left[1 + \exp\left(-\frac{x-\mu}{\sigma}\right)\right]^2} \mathbb{1}_{-\infty, \infty}(x).$$

La función de distribución asociada es

$$F(x; \mu, \sigma) = \frac{1}{[1 + \exp(-\frac{x-\mu}{\sigma})]}.$$

Para encontrar los momentos teóricos tendremos que primero jugar un poco con las funciones de densidad y distribución acumulada de esta. Primero puesto que la logística pertenece a la familia de localización y escala entonces si  $X$  es Logística  $(\mu, \sigma)$  entonces  $Y = \frac{X-\mu}{\sigma}$  se distribuye como Logística  $(0, 1)$ . Ahora veamos a la función de densidad de  $Y$

$$f_Y(y; \mu, \sigma) = \frac{e^{-y}}{(1 + e^{-y})^2},$$

y entonces veamos que si  $g(y) = yf_Y(y; \mu, \sigma)$  entonces

$$-g(y) = -yf_Y(y; \mu, \sigma) = -\frac{ye^{-y}}{(1 + e^{-y})^2} = \frac{-ye^y}{(1 + e^y)^2} = -yf_Y(-y; \mu, \sigma) = g(-y),$$

por lo tanto  $g$  es impar analogamente podemos verificar que  $yg(y)$  es par. Ahora juguemos un poco con la función de distribución acumulada de  $Y$

$$F_Y(y; \mu, \sigma) = \frac{1}{[1 + e^{-y}]},$$

luego para  $y > 0$  tenemos que

$$\frac{1}{[1 + e^{-y}]} = \sum_{n=1}^{\infty} (-1)^n e^{-yn},$$

por lo tanto

$$\frac{\partial F_Y}{\partial y} = \sum_{n=1}^{\infty} n(-1)^{n-1} e^{-yn}.$$

Ahora si estamos listos para calcular la esperanza y varianza de  $Y$ . La esperanza de  $Y$  por definición es

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty} \frac{ye^{-y}}{(1 + e^{-y})^2} dy,$$

pero al ser  $g(y)$  una función impar entonces  $\mathbb{E}[Y] = 0$  y por lo tanto  $\mathbb{E}[X] = \mu$  (linealidad de la esperanza). Mientras que la Varianza por definición es

$$Var(Y) = \int_{-\infty}^{\infty} \frac{y^2 e^{-y}}{(1 + e^{-y})^2} dy,$$

al ser  $yg(y)$  una función par tenemos que

$$Var(Y) = 2 \int_0^{\infty} \frac{y^2 e^{-y}}{(1 + e^{-y})^2} dy,$$

luego nos conviene expresar la anterior integral como

$$Var(Y) = 2 \int_0^{\infty} y^2 \frac{\partial^2 F}{\partial d \partial y},$$

por lo demostrado anteriormente tenemos que

$$Var(Y) = 2 \int_0^\infty y^2 \sum_{n=1}^\infty n(-1)^{n-1} e^{-yn} dy,$$

por convergencia dominada tenemos que

$$Var(Y) = 2 \sum_{n=1}^\infty n(-1)^{n-1} \int_0^\infty y^2 e^{-yn} dy,$$

esto es

$$Var(Y) = 2 \sum_{n=1}^\infty n(-1)^{n-1} \int_0^\infty y^2 e^{-yn} dy,$$

por integración por parte obtenemos

$$Var(Y) = 2 \sum_{n=1}^\infty n(-1)^{n-1} \left( -\frac{e^{-nx} (n^2 x^2 + 2nx + 2)}{n^3} \Big|_0^\infty \right),$$

puesto que la función exponencial  $e^{-nx}$  decrece más rapido que  $(n^2 x^2 + 2nx + 2)$  entonces esto es

$$Var(Y) = 2 \sum_{n=1}^\infty n(-1)^{n-1} \frac{2}{n^3} = 4 \sum_{n=1}^\infty (-1)^{n-1} \frac{1}{n^2} = \frac{\pi^2}{3},$$

La última igualdad se sigue del problema de Basilea. Por lo tanto  $Var(X) = \frac{\pi^2 \sigma^2}{3}$ .

En resumen tenemos

- $\mathbb{E}[X] = \mu.$
- $Var(X) = \frac{\pi^2 \sigma^2}{3}.$
- $\mathbb{E}[X^2] = \mu^2 + \frac{\pi^2 \sigma^2}{3}.$

Por lo tanto los estimadores de momentos son

- $\mu = \mathbb{E}[X] = \frac{\sum_{i=1}^n x_i}{n}$
- $\sigma = \sqrt{\frac{3Var(X)}{\pi^2}} = \sqrt{\frac{3 \sum_{i=1}^n (x_i - \bar{X})^2}{n\pi^2}}.$

c) Gaussiana Inversa  $(\mu, \lambda)$  donde la media  $\mathbb{E}[X] = \mu$  es positiva y  $\lambda$  es un parámetro de forma positivo. La varianza es  $Var(X) = \frac{\mu^3}{\lambda}$ . La densidad es

$$f(x; \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left(\frac{\lambda}{\mu}\right) \exp\left(-\frac{\lambda}{2\mu^2}x - \frac{\lambda}{2x}\right) \mathbb{1}_{(0, \infty)}(x).$$

Los momentos teóricos son

- $\mathbb{E}[X] = \mu.$
- $Var(X) = \frac{\mu^3}{\lambda}.$
- $\mathbb{E}[X^2] = \mu^2 + \frac{\mu^3}{\lambda}.$

Por lo tanto los estimadores de momentos son

- $\mu = \mathbb{E}[X] = \frac{\sum_{i=1}^n x_i}{n}$

$$\blacksquare \lambda = \frac{\mathbb{E}[X]^3}{Var(X)} = \frac{(\sum_{i=1}^n x_i)^2}{n^2 \sum_{i=1}^n (x_i - \bar{X})^3}$$

d) Gumbel  $(a, b)$  de máximos donde  $\mathbb{E}(X) = a + b\gamma$ , con  $\gamma = 0,5772$ , la varianza es  $Var(X) = \frac{\pi^2 b^2}{6}$  y la densidad es

$$f(x; a, b) = \frac{1}{b} \exp \left[ -\frac{x-a}{b} - \exp \left( -\frac{x-a}{b} \right) \right] \mathbb{1}_{(-\infty, \infty)}(x).$$

La distribución es entonces

$$F(x; a, b) = \exp \left[ -\exp \left( -\frac{x-a}{b} \right) \right].$$

Los momentos teóricos son

$$\begin{aligned} \blacksquare \mathbb{E}[X] &= a + b\gamma. \\ \blacksquare Var(X) &= \frac{\pi^2 b^2}{6}. \\ \blacksquare \mathbb{E}[X^2] &= (a + b\gamma)^2 + \frac{\pi^2 b^2}{3}. \end{aligned}$$

Por lo tanto los estimadores de momentos son

$$\begin{aligned} \blacksquare a &= \mathbb{E}[X] - \gamma \sqrt{\frac{3Var(X)}{\pi^2}} = \frac{\sum_{i=1}^n x_i}{n} - \left( \sqrt{\frac{3 \sum_{i=1}^n (x_i - \bar{X})^2}{n\pi^2}} \right) \gamma \\ \blacksquare b &= \sqrt{\frac{3Var(X)}{\pi^2}} = \sqrt{\frac{3 \sum_{i=1}^n (x_i - \bar{X})^2}{n\pi^2}}. \end{aligned}$$

e) Lognormal  $(\mu, \sigma)$  donde  $\exp(\mu)$  es la mediana y la densidad es

$$f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp \left( -\frac{\ln x - \mu}{2\sigma^2} \right) \mathbb{1}_{(0, \infty)}(x).$$

Empezaremos por ver como es la función generadora de momentos de una variable aleatoria  $U$  normal estandar

$$\begin{aligned} M_U(t) &= \mathbb{E}[e^{tU}] = \int_{-\infty}^{\infty} e^{tu} \left( \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} \right) du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tu} \left( e^{-\frac{u^2}{2}} \right) du \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(u-t)^2}{2} + \frac{t^2}{2}} du = \frac{e^{\frac{t^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(u-t)^2}{2}} du = \frac{e^{\frac{t^2}{2}}}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-v^2} dv = e^{\frac{t^2}{2}}. \end{aligned}$$

Ahora veamos que si  $X$  es Lognormal  $(\mu, \sigma)$  entonces  $X = e^{\mu + \sigma U}$  por lo que

$$\mathbb{E}[X^\alpha] = \mathbb{E}[e^{\alpha\mu + \alpha\sigma U}] = e^{\alpha\mu} \mathbb{E}[e^{\alpha\sigma U}] = e^{\alpha\mu} M_U(\alpha\sigma) = e^{\alpha\mu + \frac{\alpha^2 \sigma^2}{2}}.$$

Poe lo tanto los momentos teóricos son

$$\begin{aligned} \blacksquare \mathbb{E}[X] &= e^{\mu + \frac{\sigma^2}{2}}. \\ \blacksquare Var(X) &= e^{2(\mu + \sigma^2)} - e^{2\mu + \sigma^2}. \\ \blacksquare \mathbb{E}[X^2] &= e^{2(\mu + \sigma^2)}. \end{aligned}$$

y concluimos que los estimadores de momentos son

$$\blacksquare \mu = \frac{\log \left( \frac{\mathbb{E}[X^2]}{\mathbb{E}[X]^4} \right)}{2} = \frac{\log \left( n \frac{\sum_{i=1}^n x_i^2}{(\sum_{i=1}^n x_i)^4} \right)}{2}$$

$$\blacksquare \sigma = \sqrt{\log\left(\frac{\mathbb{E}[X^2]}{\mathbb{E}[X]^2}\right)} = \sqrt{\log\left(n \frac{\sum_{i=1}^n x_i^2}{\left(\sum_{i=1}^n x_i\right)^2}\right)}$$

4. Para la distribución normal estimada y para tu distribución asignada estimada presenta las siguientes figuras.
  - Grafica la función de distribución empírica de los datos observados y encima en la misma gráfica la distribución normal estimada y tu distribución estimada.
  - Grafica la densidad normal estimada y la densidad estimada de tu distribución asignada en la misma figura. En el eje horizontal a una altura de cero grafica con asteriscos los datos observados.
  - Gráfica *PP* para el modelo normal usando los datos observados.
  - Gráfica *PP* para tu distribución asignada para los datos observados.
5. Con base a todos tus resultados anteriores, comenta si el modelo normal es razonable para los datos de nuestro grupo o si tu modelo asignado sería una mejor opción? Grafica la densidad estimada con el que decidas es mejor modelo contra la densidad normal estimada para los datos anteriores de alumnos del pasado de *CIMAT* y *DEMAT*. Compara y comenta.