Universidad del Valle de Guatemala
Departamento de Matemática
Licenciatura en Matemática Aplicada
Fecha de entrega: 21 de marzo de 2021
Rudik R. Rompich - Carné: 19857

Estadística Matemática - Paulo Mejía

Parcial 2

Instrucciones: Resuelva los siguientes problema. Favor hacer la solución en latex y cargar el archivo latex y pdf en la tarea de Canvas.

Para la resolución de este parcial, se usarán las definiciones y teoremas de Wackerly et al. (2014).

1. Problema 1

Sean Y_1 y Y_2 variables aleatorias continuas que tiene una función densidad conjunta:

$$f(y_1, y_2) = \begin{cases} e^{-(y_1+y_2)} & , \quad 0 \leq y_1, 0 \leq y_2 \\ 0 & , \quad \text{en cualquier otro caso} \end{cases}$$

1. a) Encontrar la función densidad de $U = Y_1 + Y_2$.

   Solución. Debido al método usado para el desarrollo del problema, primero se realizó el inciso **2. b)**. Simplemente consiste en derivar la función encontrada es decir:

$$f_U(u) = \frac{dF_U(u)}{du} \tag{1}$$

$$= \begin{cases} ue^{-u} & u \geq 0 \\ 0 & u < 0 \end{cases} \tag{2}$$

■

2. b) Calcular la función de distribución acumulada de U.

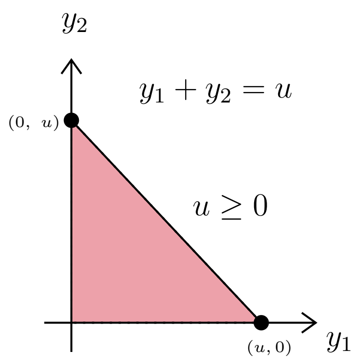
   Solución. Para el desarrollo del problema se plantea el método de las funciones de distribución:

Summary of the Distribution Function Method

Let U be a function of the random variables $Y_1, Y_2, \ldots, Y_n$.

1. Find the region $U = u$ in the $(y_1, y_2, \ldots, y_n)$ space.
2. Find the region $U \leq u$.
3. Find $F_U(u) = P(U \leq u)$ by integrating $f(y_1, y_2, \ldots, y_n)$ over the region $U \leq u$.
4. Find the density function $f_U(u)$ by differentiating $F_U(u)$. Thus, $f_U(u) = dF_U(u)/du$.

Primero, se encuentra la región $U = u$, es decir, tenemos $y_1 + y_2 = u$:



Segundo, encontramos la región en donde $U \leq u$. Por la figura, notamos que la región es la parte inferior a la recta u. Si $u < 0$, entonces $F_U(u) = 0$. Por otra parte, si $u \geq 0$, tenemos la región en rojo.

Procedemos a encontrar la región la región en rojo delimitado por $(0,0), (0,u), (u,0)$, es decir que tenemos:

$$\begin{aligned}
F_U(u) &= \int_0^u \int_0^{-y_1+u} f(y_1, y_2)\, dy_2 dy_1 && (1)\\
&= \int_0^u \int_0^{-y_1+u} e^{-(y_1+y_2)} dy_2 dy_1 && (2)\\
&= -ue^{-u} - e^{-u} + 1 && (3)
\end{aligned}$$

Es decir, que la función de distribución acumulada es:

$$F_U(u) = \begin{cases} -ue^{-u} - e^{-u} + 1 & u \geq 0 \\ 0 \quad u < 0 \end{cases} \tag{4}$$

■

3. c) Determine el valor $E(U)$.

Solución.

$$\begin{aligned}
E(U) &= \int_{-\infty}^{\infty} u f_U(u) du \\
&= \int_0^{\infty} u \left(ue^{-u}\right) du \\
&= \int_0^{\infty} u^2 e^{-u} du \\
&= -u^2 e^{-u} - 2ue^{-u} - 2e^{-u}\Big|_0^{\infty} \\
&= 2
\end{aligned} \tag{1}$$

■

(Valor 25 puntos).

2. Problema 2

Sean Y_1 y Y_2 variables aleatorias de Poisson independientes con medias λ_1 y λ_2

1. a) Encontrar la función generadora de momentos de la variable aleatoria Y_1.

   Solución. En el Apéndice 2 - Cuadro 1, se ofrece la función generadora de momentos para Poisson.

   Table 1 Discrete Distributions

Distribution	Probability Function	Mean	Variance	Moment-Generating Function
Poisson	$p(y) = \dfrac{\lambda^y e^{-\lambda}}{y!}$; $y = 0, 1, 2, \dots$	λ	λ	$\exp[\lambda(e^t - 1)]$

   La deducción para *el caso general* es la siguiente:

$$
\begin{aligned}
m_U(t) &= E(e^{tU}) && && (1)\\
&= \sum_{U=0}^{\infty} e^{tU} \cdot p(U) && \text{Definición 5.9} && (2)\\
&= \sum_{U=0}^{\infty} e^{tU} \cdot \frac{\lambda^U e^{-\lambda}}{U!} && U = 0, 1, 2, ... && (3)\\
&= e^{-\lambda} \sum_{U=0}^{\infty} \frac{e^{tU}\lambda^U}{U!} = e^{-\lambda} \sum_{U=0}^{\infty} \frac{(e^t\lambda)^U}{U!} && && (4)
\end{aligned}
$$

   Se observa que se tiene una serie de una función exponencial, i.e. $\sum_{k=0}^{\infty} \frac{z^k}{k!} = e^z$.

$$
= e^{-\lambda} e^{e^t\lambda} = e^{-\lambda + e^t\lambda} = e^{\lambda(e^t-1)} \tag{5}
$$

   Expresando la solución en una notación más clara:

$$
= \exp[\lambda(e^t - 1)] \tag{6}
$$

   Entonces, para el caso específico de Y_1, se tiene:

$$
m_{Y_1}(t) = \exp[\lambda_1(e^t - 1)] \tag{7}
$$

   Analogámente para Y_2:

$$
m_{Y_2}(t) = \exp[\lambda_2(e^t - 1)] \tag{8}
$$

$\blacksquare$

2. b) Determinar la función generadora de momentos de la variable aleatoria $Y_1 + Y_2$.

Solución. Tomando como referencia el Teorema 6.2:

> Let $Y_1, Y_2, \ldots, Y_n$ be independent random variables with moment-generating functions $m_{Y_1}(t), m_{Y_2}(t), \ldots, m_{Y_n}(t)$, respectively. If $U = Y_1 + Y_2 + \cdots + Y_n$, then
> $$m_U(t) = m_{Y_1}(t) \times m_{Y_2}(t) \times \cdots \times m_{Y_n}(t).$$

$$\begin{aligned} m_{Y_1+Y_2} &= m_{Y_1}(t) \times m_{Y_2}(t) && (1)\\ &= \exp[\lambda_1(e^t - 1)] \times \exp[\lambda_2(e^t - 1)] && (2)\\ &= \exp[\lambda_1(e^t - 1) + \lambda_2(e^t - 1)] && (3) \end{aligned}$$

Entonces, tenemos la función generadora de momentos $(Y_1 + Y_2)$ con media $\lambda_1 + \lambda_2$:

$$= \exp[(\lambda_1 + \lambda_2)(e^t - 1)] \tag{4}$$

■

3. c) Encontrar la función de probabilidad de $Y_1 + Y_2$, utilizando los incisos b) y a).

Solución. Es decir, usando nuevamente el Cuadro 1 - Apéndice 2, se necesita regresar la función generadora de momentos a su expresión original:

$$m_{Y_1+Y_2} = \exp[(\lambda_1 + \lambda_2)(e^t - 1)]$$

Es decir, que tenemos:

$$P(Y_1 + Y_2) = \frac{(\lambda_1 + \lambda_2)^{Y_1+Y_2} e^{-(\lambda_1+\lambda_2)}}{(Y_1 + Y_2)!}$$

■

4. d) Determinar la función de probabilidad condicional de Y_1, dado $Y_1 + Y_2 = m$.

Solución. Como es ampliamente conocido, la probabilidad condicional de una variable A dado una variable B , se calcula con $P(A|B) = \frac{P(A \cap B)}{P(B)}$. Además, como sabemos que el problema trata con variables independientes, entonces la ecuación queda como: $P(A|B) = \frac{P(A)P(B)}{P(B)}$. Usaremos la notación usual, Y_n es la variable aleatoria y y_n es un valor particular de Y_n. Entonces:

$$P(Y_1 = y_1 | Y_1 + Y_2 = m) = \frac{P(Y_1 = y_1)P(Y_1 + Y_2 = m)}{P(Y_1 + Y_2 = m)} \tag{1}$$

Como tenemos que encontrar la probabilidad condicional de Y_1 dado $Y_1 + Y_2 = m$ y desconocemos la probabilidad de Y_2, por lo que es necesario despejar Y_2:

$$= \frac{P(Y_1 = y_1)P(Y_2 = m - y_1)}{P(Y_1 + Y_2 = m)} \tag{2}$$

Considerando la función probabilidad de Poisson:

$$= \frac{\frac{\lambda_1^{y_1} e^{-\lambda_1}}{y_1!}\left(\frac{\lambda_2^{m-y_1} e^{-\lambda_2}}{(m-y_1)!}\right)}{\frac{(\lambda_1+\lambda_2)^m e^{-(\lambda_1+\lambda_2)}}{m!}} \tag{3}$$

Luego de hacer el desarrollo algebraico, notamos que la distribución binomial, específicamente la función de masa de probabilidad, aparece. $P_k = \binom{n}{k} p^k (1-p)^{n-k}$, en donde $\binom{n}{k} = \frac{n!}{k!(n-k)!}$, por lo que el resultado es:

$$= \begin{pmatrix} m \\ y_1 \end{pmatrix} \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{y_1} \left(1 - \frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{m-y_1} \tag{4}$$

■

(Valor 25 puntos).

3. Problema 3

Los tiempos de servicio para los clientes que pasan por una caja de un supermercado son variables aleatorias independientes con $\mu = 1{,}5$ minutos y $\sigma^2 = 1{,}0$. Se quiere calcular la probabilidad de que 100 clientes puedan ser atendidos en menos de 2 horas del tiempo total del servicio.

1. a) ¿Se puede aplicar el teorema del límite central?

Solución. Considerando el teorema de límite central:

Central Limit Theorem: Let $Y_1, Y_2, \ldots, Y_n$ be independent and identically distributed random variables with $E(Y_i) = \mu$ and $V(Y_i) = \sigma^2 < \infty$. Define

$$U_n = \frac{\sum_{i=1}^n Y_i - n\mu}{\sigma\sqrt{n}} = \frac{\overline{Y} - \mu}{\sigma/\sqrt{n}} \qquad \text{where } \overline{Y} = \frac{1}{n}\sum_{i=1}^n Y_i.$$

Then the distribution function of U_n converges to the standard normal distribution function as $n \to \infty$. That is,

$$\lim_{n\to\infty} P(U_n \le u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt \qquad \text{for all } u.$$

Se sabe que si una muestra tiene más de 30 elementos, entonces se asegura que se podrá tratar aproximadamente como una distribución normal. En un contexto formal, para una distribución con un $n \to \infty$ (n lo suficiente grande) entonces la probabilidad podrá ser aproximada por una distribución normal.

Como nos lo asegura el teorema, ya que tenemos $\mu = 1{,}5$, $\sigma^2 = 1$ y muestras $Y_1, Y_2, ..., Y_{100}$. (En donde $n > 30$, que es significativo). Lo que implica que se puede definir un U_n. Por lo tanto, el problema cumple las condiciones adecuadas para aplicar el Teorema del Límite Central. ■

2. b) Calcule la probabilidad de que 100 clientes puedan ser atendidos en menos de 2 horas del tiempo total del servicio (Use R).

   Solución. Comenzamos colocando todos los datos en una misma dimensional, es decir: Media = 1,5 minutos, Varianza = 1 minuto, Tiempo total de servicio = 120 minutos. Entonces, nos piden la probabilidad que 100 clientes puedan ser atendidos en 120 minutos, es decir, usando un *script* de R:

```
clientes <- 100
tiempo <- 120
media <-1.5
dev_est <- 1/sqrt(100)
pnorm(tiempo/clientes,media, dev_est)
```

```
## [1] 0.001349898
```

■

3. c) ¿Considera que es posible atender a los 100 clientes en menos de 2 horas?

   Solución. No, ya que la probabilidad es ínfima; insignificante. Aunque puede afirmarse que no es imposible, solo que bastante improbable. ■

(Valor 25 puntos).

4. Problema 4

Sean $Y_1, Y_2, \ldots, Y_n$ variables aleatorias normales e independientes, cada una con media μ y varianza σ^2.

1. a) Determine la distribución de $\sqrt{n}\left(\frac{\bar{Y}-\mu}{\sigma}\right)$ (no tiene que encontrar la forma cerrada de la distribución sólo indicar el tipo con sus parámetros)

   Solución. Es trivial asumir que $\mu = \mu_{\bar{Y}}$ y $\sigma = \sqrt{\sigma_{\bar{Y}}^2 n} = \sigma_{\bar{Y}}\sqrt{n}$. Entonces por el Teorema 7.1

   > Let $Y_1, Y_2, \ldots, Y_n$ be a random sample of size n from a normal distribution with mean μ and variance σ^2. Then
   > $$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$
   > is normally distributed with mean $\mu_{\overline{Y}} = \mu$ and variance $\sigma_{\overline{Y}}^2 = \sigma^2/n$.

   es posible asumir que tiene una distribución normal estándar. En donde: σ = desviación estándar, n= número de variables, $\bar{Y}$= función de media muestral de variables aleatorias y μ = media. ■

2. b) Determine la distribución de $\frac{(n-1)S^2}{\sigma^2}$ (no tiene que encontrar la forma cerrada de la distribución sólo indicar el tipo con sus parámetros)

   Solución. Por el Teorema 7.3

> Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from a normal distribution with mean μ and variance σ^2. Then
>
> $$\frac{(n-1)S^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(Y_i - \overline{Y})^2$$
>
> has a χ^2 distribution with $(n-1)$ df. Also, $\overline{Y}$ and S^2 are independent random variables.

se puede determinar que la distribución tiene una distribución χ^2 con $(n-1)$ grados de libertad. En donde: n número de variables, S^2 varianza muestral y σ^2 la varianza. ■

3. c) Determine la distribución de $\sqrt{n}\left(\frac{\bar{Y}-\mu}{S}\right)$ (no tiene que encontrar la forma cerrada de la distribución sólo indicar el tipo con sus parámetros)

   Solución. Es un caso especial del problema **1. a)** ya que si σ es desconocido, entonces tiene que ser estimado por $S = \sqrt{S^2}$: Por lo que es posible asumir que tiene una distribución t de Student. En donde: $S =$ desviación estándar muestral, $n=$ número de variables, $\bar{Y}=$ función de media muestral de variables aleatorias y $\mu =$ media. ■

4. d) Sean 2 muestras aleatorias independientes tomadas de distribuciones normales de tamaño n_1 y n_2. Determine la distribución de $\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$ (no tiene que encontrar la forma cerrada de la distribución sólo indicar el tipo con sus parámetros)

   Solución. De acuerdo a la definición 7.3:

   > Let W_1 and W_2 be *independent* χ^2-distributed random variables with ν_1 and ν_2 df, respectively. Then
   >
   > $$F = \frac{W_1/\nu_1}{W_2/\nu_2}$$
   >
   > is said to have an F distribution with ν_1 numerator degrees of freedom and ν_2 denominator degrees of freedom.

   por la forma de la distribución, entonces esta debe tener una distribución F con $n_1 - 1$ grados de libertad en el numerador y $n_2 - 1$ en el denominador. En donde: S_1 y S_2 representan la desviación estándar muestral y σ_1^2 y σ_2^2 representan la varianza. ■

(Valor 25 puntos).

Referencias

Wackerly, D., Mendenhall, W., and Scheaffer, R. L. (2014). *Mathematical statistics with applications*. Cengage Learning.