Tarea 1

Análisis Estadístico Multivariado

Emmanuel Alcalá

Ejercicio 1

Considera la distribución bivariada $p(x, y)$ de dos VA discretas X, Y

	y_1	0.01	0.02	0.03	0.1	0.1
Y	y_2	0.05	0.1	0.05	0.07	0.2
	y_3	0.1	0.05	0.03	0.05	0.04
		x_1	x_2	x_3	x_4	x_5
				X		

Figure 1: Distribución conjunta $p(y, x)$

Calcula:

1. Las distribuciones marginales $p(x)$ y $p(y)$
2. Las distribuciones condicionales $p(x \mid Y = y_1)$ y $p(y \mid X = x_3)$

Ejercicio 2

Supongamos lo siguiente:

Un 5% de la población son mujeres (M) con estudios universitarios completos (U). Las mujeres representan un 55% de la población. Un 20% de la población tiene estudios universitarios completos

1. ¿Cuál es la probabilidad de que una mujer tenga estudios universitarios completos? Hint: calcular $p(U \mid M)$.
2. ¿Cuál es la probabilidad de que una mujer tenga estudios universitarios completos? Hint: calcular $p(M \min U)$
3. ¿Cuál es la probabilidad de que una persona con estudios universitarios completos sea hombre (H)? Hint: necesitamos el complemento de 2).

Ejercicio 3

La pmf de una distribución de probabilidad discreta de Poisson permite calcular la probabilidad un número dado de eventos en un intervalo, dada una tasa de ocurrencia λ. Por ejemplo, ¿cuál es la probabilidad de que ocurran 3 choques en un día en López Mateos si la tasa de ocurrencia es de 2 choques por día? Con la pmf

$$p(x \mid \lambda) = \frac{e^{-\lambda}\lambda^x}{x!} \tag{11}$$

Sería

$$p(3 \mid 2) = \frac{e^{-2}\lambda^3}{3!}.$$

En `R`, la probabilidad de que x tome el valor de 3 si el valor de λ es 2 se calcularía así (creamos nuestra propia función a partir de la ecuación (11)):

```
# creamos una función para la eq 11; esta función necesita dos argumentos:
# x y lambda
pmf_poisson <- function(x, lambda) (exp(-lambda) * lambda^x) / (factorial(x))
# asignamos valores a los argumentos
x <- 3
l <- 2 # lambda
# usamos la función creada ingresando los argumentos
pmf_poisson(x = x, lambda = l)
```

```
[1] 0.180447
```

```
# NOTA: R tiene una función base para hacer esto, dpois(),
# pero estos ejercicios son buenos para desmitificar R mientras nos hábiles
# dpois(x, l) # da el mismo resultado
```

Gráficamente, nuestro valor $x = 3$ en una distribución de Poisson se vería así:

```
par(las=1)
xp <- seq(0, 15, 1)
# graficar la distribución (en barras)
plot(
  xp,
  pmf_poisson(xp, l),
  type = "h",
  ylab = "P(X=x)",
  xlab = 'x',
  main = expression(Poisson~pmf~p(x*'|'*lambda==2))
)
# enfatizar la barra de probabilidad de x=3
segments(
  # inicio del segmento en x
  x0 = 3,
  # fin del segmento en x
  x1 = 3,
  # inicio del segmento en y
  y0 = 0,
  # altura del segmento
  y1 = pmf_poisson(x = x, l = l),
  col = 'red',
  lwd = 2
)
```

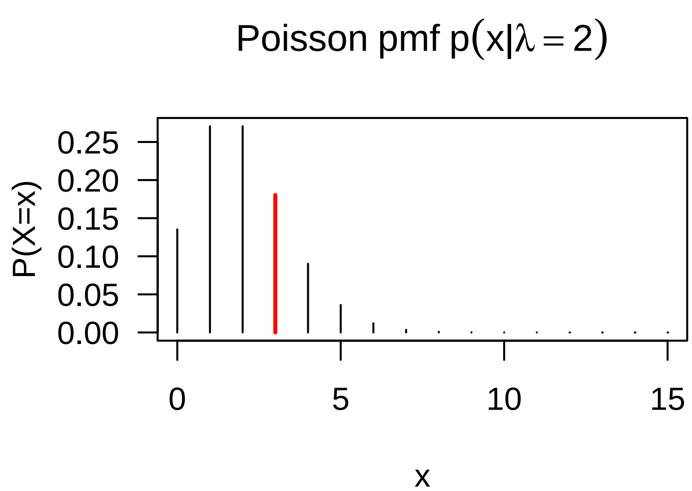

Figure 2: Distribución de probabilidad de Poisson con $x = 3$ y $\lambda = 2$

Resolver (y graficar):

1. $p(x = 3 \mid \lambda = 3/4) = ?$
2. $p(0 \leq x \leq 10 \mid \lambda = 3) = ?$

Ejercicio 4

La pdf de la distribución normal con media μ y desviación estándar σ (o varianza σ^2) es caracterizada por

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \tag{12}$$

1. Suponer que un manufacturador de un tipo de botana sabe que el peso total del paquete de botana está distribuido normalmente con una media de 80.2 gramos y una desviación estándar de 1.1 gramos. ¿Cuál es la probabilidad de que un paquete de botana pese *menos* de 78 gramos?
2. Bajo las mismas asunciones anteriores, ¿cuál es la probabilidad de que un paquete dado tenga un peso que esté entre 2 desviaciones estándar de la media? Es decir, $p(x \in [\mu - \sigma, \mu + \sigma])$.