

Econometría I

Reto 02

Carlos A. Yanes Guerra

Universidad del Norte | Departamento de Economía

Contenido

Antes de empezar	1
Objetivo	1
Base a usar	1
Caso 1	3
Caso 2	3
Caso 3	3
Caso especial (No Obligatorio)	4

Antes de empezar

Recuerde que los retos son para desarrollarlos en clases en el tiempo sugerido por el profesor **1H:20 Min.** Trate con su grupo de trabajo (máximo 2 personas) este *-tenga un toque único-* donde haga uso en la gran mayoría de los códigos de clase y no los que les da la IA.

Objetivo

Aplicar el análisis de distribuciones de probabilidad

Base a usar

Corra el siguiente código:

```
# Para la base de datos
usuarios <- data.frame(
  Usuario_ID = 1:10,
  Edad = c(25, 34, 29, 42, 37, 22, 31, 28, 45, 26),
  Género = c("M", "F", "M", "F", "M", "F", "M", "M", "F", "F"),
  Servicio_Favorito = c(3, 2, 1, 5, 4, 3, 2, 1, 5, 4),
  Horas_Semanales = c(10, 5, 15, 8, 7, 12, 9, 6, 4, 11),
  Satisfacción = c(4, 3, 5, 2, 3, 4, 5, 3, 2, 4)
)

# Una manera de convertir las variables categóricas en factores (cualitativas)
usuarios$Género <- factor(usuarios$Género, levels = c("M", "F"))
usuarios$Servicio_Favorito <- factor(usuarios$Servicio_Favorito,
  levels = c(1, 2, 3, 4, 5),
  labels = c("Direct Tv", "Amazon Prime", "Netflix", "iTunes", "Tv Cable"))
```

```
# Mostrar la base de datos  
head(usuarios)
```

Caso 1

Saber de las distribuciones, nos permite resolver **problemas** que se presentan a continuación: Sea X un conjunto de balotas de color verde que se pueden extraer de una bolsa que regularmente solo contiene un 15% de ese color. Con los siguientes parámetros como $n=20$ total de pelotas dentro de la bolsa y $p = 15\% = 0.15$. Si queremos conocer la **probabilidad** de sacar 3 pelotas de color verde, debemos entonces hacer uso de una distribución. Todos sabemos que la probabilidad va en $x \in \{0, 1, 2, \dots, 20\}$, la propuesta de desarrollo es:

$$f(x) = P(X = x) = \binom{n}{x} \cdot P^x \cdot (1 - P)^{n-x} = \binom{20}{x} \cdot 0.15 \cdot (0.85)^{20-x}$$

La pregunta va en dirección de la probabilidad de sacar 3 balotas de color verde, lo que es:

$$f(x) = \binom{20}{3} \cdot 0.15 \cdot (0.85)^{20-3}$$

En el Software R es muy fácil aplicar el comando `dbinom(x,n,p)` y tendremos:

```
dbinom(3,20,0.15)
```

Dándonos como resultado una probabilidad de 24.28% de sacar 3 pelotas de color verde.

Preguntas

- Haga una secuencia con 20 observaciones, genere un gráfico de distribución binomial tipo discreto.
- Cuál es la probabilidad de sacar 5 balotas verdes pero cuando n es 240 y la probabilidad de ese color dentro de la bolsa sea de 30%?

Caso 2

Supongamos que x es una variable aleatoria que sigue una distribución normal con media $\mu = 3$ y $\sigma^2 = 4$

Desarrolle lo siguiente

- Calcule la $P(X > 5)$
- Calcule $P(2 < X < 6)$

Caso 3

El archivo de datos proporcionado al inicio de esta guía contiene información de 10 estudiantes de sus hábitos televisivos, además de las horas que lo hace en la semana, su edad y por ultimo el genero de la persona. De acuerdo a esa base de datos proporcionada, responda lo siguiente (ojo! utilice los comandos de las guías del curso)

- ¿Qué porcentaje de hombres y mujeres ven solo Amazon Prime?
- ¿Considera usted que la edad y el consumo del entretenimiento tienen alguna relación?
- ¿Qué tipo de servicio de televisión es el mas usado? Cuál da mas satisfacción?
- ¿Qué tipo de distribución, curtosis y asimetría posee las horas de entretenimiento de los estudiantes?
- De la red que menor uso tiene, que le recomendaría para aumentar el consumo de entretenimiento?. *No requiere de código alguno para responder esto*

Caso especial (No Obligatorio)

Piense que x es el tiempo de vuelo de un avión que viaja de Lima (perú) a Bogotá (Colombia). Además tiene información que el tiempo de vuelo mínimo es 2 horas y que el tiempo máximo es 2 horas 20 minutos; por tanto, en función de los minutos, el tiempo de vuelo puede ser cualquier valor en el intervalo de 120 a 140 minutos (por ejemplo, 124 minutos, 125.48 minutos, etc.) Como la variable **aleatoria** x puede tomar cualquier valor de 120 a 140 minutos, x es una variable aleatoria continua en vez de ser discreta. Imagine que contamos con datos sobre el vuelo real suficientes para concluir que la probabilidad de un tiempo de vuelo entre 120 y 121 minutos es la misma que la probabilidad del tiempo de vuelo dentro de cualquier otro intervalo de 1 a 140 minutos. Como cada intervalo de 1 minuto es igualmente probable, la variable aleatoria x tiene una distribución de probabilidad uniforme.

La función de probabilidad viene a ser:

$$f(x) \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{si es lo contrario} \end{cases}$$

Donde $a = 120$ y $b = 140$

Preguntas

- ¿Cuál es la probabilidad de que el vuelo se demore entre 120 y 150 minutos?
- ¿Cuál considera usted que es la probabilidad que el vuelo dure menos de 100 minutos?.
- Determine la probabilidad que el vuelo dure mas de 135 minutos.