



EJERCICIO PRÁCTICO 3: SIMULANDO DISTRIBUCIONES EN R

CONTEXTO

Hemos estudiado las distribuciones de probabilidad por bastante tiempo. Trataremos de verlas en la práctica, con la ayuda de R.

OBJETIVOS DE APRENDIZAJE

1. Comprender la relación entre las distribuciones de probabilidad, tanto continuas como discretas.
2. Practicar el manejo de datos, funciones y gráficos en el entorno R.

ÉXITO DE LA ACTIVIDAD

1. El equipo consigue construir las distribuciones solicitadas utilizando las definiciones estudiadas, no las funciones disponibles de R `dxxx()`, `pxxx()`, `qxxx()` y `rxxx()`.
2. El equipo consigue graficar las distribuciones obtenidas para compartirlas con sus compañeras y compañeros.

ACTIVIDADES

El siguiente código permite obtener (sintéticamente) una distribución normal con la media y desviación estándar vistos en los ingresos totales registrados por la encuesta Casen 2017.

```
basename <- "Casen 2017.csv"
file <- file.path(dir, basename)
población <- read.csv(file = file)
tamaño <- nrow(población)
ingreso <- as.numeric(población[["ytot"]])
poda <- 0.2
q20 <- quantile(ingreso, poda)
q80 <- quantile(ingreso, 1 - poda)
ingreso.podado <- ingreso[ingreso > q20 & ingreso < q80]
tamaño.podado <- length(ingreso.podado)
media.ingreso <- mean(ingreso.podado)
sd.ingreso <- sqrt(sum((ingreso.podado - media.ingreso)^2) / tamaño.podado)
set.seed(133)
ingreso.normal <- rnorm(5000, mean = media.ingreso, sd = sd.ingreso)
```

Usando este código ejemplo:

1. Creen un script en R con comentarios iniciales que indiquen el número de grupo y los nombres de sus integrantes (el equipo queda en libertad de excluir a aquellos integrantes que no hayan participado del desarrollo de la actividad).
2. Definan su propia semilla y obtengan 5.000 casos para una distribución de ingresos aproximadamente normal.

3. A partir de la distribución conseguida, y sin usar nuevamente la función `rnorm()`, generen la correspondiente distribución Z.
4. Con la distribución Z obtenida en el punto anterior, y sin utilizar funciones como `rchisq()`, construyan dos distribuciones χ^2 , cada una con más de 3 y menos de 15 grados de libertad.
5. Usando las dos distribuciones χ^2 generadas en el punto anterior, construyan una distribución F.
6. En cada caso, construyan un gráfico para mostrar las distribuciones obtenidas a sus compañeras y compañeros.

El siguiente código R muestra la obtención de 20 repeticiones de un ensayo de Bernoulli con éxito si se elige la encuesta de una mujer de entre los datos seleccionados desde la encuesta Casen 2017.

```
set.seed(133)
n.repeticiones <- 20

ensayo <- function(x)
  ifelse(sample(población[["sexo"]], 1) == "Mujer", 1, 0)

veinte.repeticiones <- sapply(1:n.repeticiones, ensayo)
```

Usando este código ejemplo:

7. Definan su propia semilla y número de repeticiones para el ensayo.
8. Generen, sin utilizar funciones como `rbinom()`, una distribución binomial.
9. Similarmente, construyan una distribución geométrica.
10. Análogamente, generen una distribución binomial negativa.
11. En cada caso, construyan un gráfico para mostrar las distribuciones obtenidas a sus compañeras y compañeros.

Fuera del horario de clases, un integrante de cada equipo debe subir el script realizado al espacio destinado para ello en UVirtual con el nombre "EP03-respuesta-grupo-i", donde i es el número de grupo asignado. Las respuestas deben subirse antes de las 23:30 del sábado.

CRITERIOS DE EVALUACIÓN

- Escriben código R que carga el archivo CSV y asegura que los tipos de variables se reflejen en el data frame resultante.
- Obtienen la muestras solicitadas correctamente usando una semilla propia, distinta a la usada en el ejemplo.
- Construyen correctamente una distribución Z, a partir de la distribución normal indicada en el enunciado, usando para ello operaciones y operandos adecuados.
- Construyen correctamente dos distribuciones chi-cuadrado con más de 3 y menos de 15 grados de libertad cada una, a partir de la distribución Z obtenida anteriormente.
- Construyen correctamente una distribución F, a partir de las dos distribuciones chi-cuadrado obtenidas anteriormente, usando para ello operaciones y operadores adecuados.
- Construyen correctamente una distribución binomial, a partir de la distribución de Bernoulli indicada en el enunciado, usando para ello operaciones y operandos adecuados.

- Construyen correctamente una distribución geométrica, a partir de la distribución de Bernoulli indicada en el enunciado, usando para ello operaciones y operandos adecuados
- Construyen correctamente una distribución binomial negativa, a partir de la distribución de Bernoulli indicada en el enunciado, usando para ello operaciones y operandos adecuados
- El script R es ordenado, está bien indentado, no contiene sentencias espurias y está bien comentado
- Los comentarios están escritos con buena ortografía (<5 errores) y redacción, usando vocabulario propio de la disciplina y el contexto del problema