Práctica 5

Modelos de distribuciones

5.1. Introducción

En los menús `Distribuciones`→`Distibuciones continuas/Distribuciones discretas` encontramos un conjunto de opciones que permiten representar las funciones de densidad y de probabilidad de diversos modelos de distribuciones continuas y discretas, así como la correspondiente función de distribución. También es posible calcular probabilidades, cuantiles y generar valores aleatorios de acuerdo a estos modelos.

Para los **modelos discretos** que aparecen en R-Commander disponemos de las siguientes opciones:

- `Cuantiles (modelo)`: dada una variable aleatoria X discreta, el cuantil de orden k (opción cola izquierda) es el menor valor c_k de la variable para el cual se verifica que $P(X \leq c_k) \geq k$.
- `Probabilidades (modelo) acumuladas`: función de distribución del modelo considerado. Para cada valor x la opción `cola izquierda` muestra $P[X \leq x]$. La opción `cola derecha` muestra $P[X > x]$.
- `Probabilidades (modelo)`: muestra la probabilidad de los valores que toma la variable.
- `Gráfica de la distribución (modelo)`: representa la función de probabilidad o de distribución del modelo considerado.
- `Muestra de una distribución (modelo)`: genera valores aleatorios de acuerdo al modelo considerado.

Para las **distribuciones continuas** disponemos de las siguientes opciones:

- `Cuantiles (modelo)`: el cuantil k, $0 \leq k \leq 1$, de una variable aleatoria X continua es el valor c_k para el cual $P(X \leq c_k) = k$. Nótese que $P(X \geq c_k) = 1-k$.

- `Probabilidades (modelo) acumuladas`: función de distribución del modelo considerado. La opción `cola izquierda` del menú proporciona, para cada valor x, el valor de $P[X \leq x] = P[X < x]$. La opción `cola derecha` proporciona $P[X > x] = P[X \geq x] = 1 - P[X \leq x]$.
- `Gráfica de la distribución (modelo)`: representa gráficamente la función de densidad o de distribución del modelo considerado.
- `Muestra de una distribución (modelo)`: genera valores aleatorios de acuerdo al modelo considerado.

5.2. Modelo binomial

Un examen tipo test consta de 21 preguntas, cada una de las cuáles tiene 4 posibles respuestas. Si un estudiante responde a todas las preguntas al azar,

a) Determinar la distribución del número de preguntas respondidas correctamente por el estudiante.

b) Representar las funciones de probabilidad y de distribución de esta variable.

c) Calcular la probabilidad de que el estudiante responda correctamente a,exactamente, 6 preguntas. (Sol: 0.1770398).

d) Calcular la probabilidad de aprobar el examen si para ello es necesario responder correctamente a, al menos, 11 preguntas. (Sol: 0.00642271).

5.3. Distribución de Poisson

El número de accesos por minuto a una página web sigue una distribución de Poisson de media 5.

a) Representar las funciones de probabilidad y de distribución de esta variable.

b) Determinar la probabilidad de que, en un minuto, se produzcan exactamente 4 accesos. (Sol: 0.1754673698).

c) Determinar la probabilidad de que, en dos minutos, el número de accesos sea, a lo más, de 7. (Sol: 0.2202206).

d) Nos preguntan por el número de accesos que, como máximo, se producirán en el próximo minuto. Determinar el menor valor que debemos dar como respuesta si deseamos acertar con probabilidad no inferior a 0.9. (Sol: 8 accesos).

5.4. Distribuciones geométrica y binomial negativa

Representar la función de probabilidad de una distribución geométrica de parámetro 0.25. Deducir la relación que existe entre la definición, vista en teoría, de la distribución geométrica y la definición que utiliza `R`.

Consideremos ahora el modelo de examen descrito en la Sección 5.2 (modelo binomial) y supongamos de nuevo que el alumno responde al azar a todas las preguntas. Responder a las siguientes cuestiones:

a) Determinar la probabilidad de que el alumno dé su primera respuesta correcta, en la cuarta pregunta que responde. (Sol: 0.1054687500).

b) Determinar la probabilidad de que necesite responder, como mucho, a 6 preguntas hasta dar una respuesta correcta. (Sol: 0.8220215).

c) Determinar la probabilidad de que el alumno dé su tercera respuesta correcta en la sexta pregunta que responde. (Sol: 0.0659179688).

Nota: los parámetros que se deben introducir en R para hacer los calculos correspondientes a la distribución binomial negativa, $B^-(r,p)$, son:

- *Número de éxitos*: número de éxitos que deben alcanzarse en la secuencia de experimentos de Bernouilli. Corresponde con el parámetro r de la distribución.
- *Probabilidad de éxito*: es la probabilidad de éxito en el experimento de Bernouilli asociado al experimento aleatorio. Corresponde con el parámetro p.
- *Valor de la variable*: al solicitar probabilidades de una variable $X \sim B^-(r,p)$, hay que tener en cuenta que R define $P(X = k)$, $k \in \{0, 1, 2, \ldots\}$, como la probabilidad de que se produzcan k fracasos antes del r-ésimo éxito.

5.5. Distribución hipergeométrica

Un lote de 20 dispositivos electrónicos contiene cinco dispositivos defectuosos. Si se eligen 10 dispositivos al azar, ¿cuál es la probabilidad de que 2 de los 10 dispositivos estén defectuosos? (Sol: 0.34829721).

Nota: si X sigue una distribución hipergeométrica, R define $P(X = t)$ como la probabilidad de encontrar t *bolas blancas* entre todas las bolas elegidas al azar.

5.6. Distribuciones exponencial y gamma

El tiempo de funcionamiento sin fallos de un ordenador sigue una variable aleatoria con distribución exponencial de media 20 minutos.

a) Representar las funciones de densidad y de distribución de esta variable.

b) Determinar la probabilidad de que el tiempo de funcionamiento sin fallos sea, al menos, 22 minutos. (Sol: 0.3328711).

c) Determinar un periodo de tiempo tal que, con probabilidad 0.99, podamos afirmar que el ordenador trabajará sin fallos durante ese periodo. (Sol: 0.2010067 min.).

d) Determinar la probabilidad de que el tiempo entre cuatro fallos consecutivos sea inferior a 60 minutos. (Sol: 0.5768099).

Nota: si $X \sim G(a, p)$, entonces el parámetro de forma al que hace referencia R es p y el parámetro de escala es $1/a$.

5.7. Distribución normal

El contenido de zumo en litros, que un proceso de llenado automático deposita en las botellas, sigue una distribución normal de media 2 y desviación típica 0.1.

a) Representar las funciones de densidad y de distribución de esta variable.

b) Determinar la probabilidad de que una botella elegida al azar contenga más de 1.9 litros de zumo. (Sol: 0.8413447).

c) Determinar la probabilidad de que una botella elegida al azar contenga entre 1.95 y 2.1 litros de zumo. (Sol: 0.5328072).

d) El llenado de una botella se considerará defectuoso si la cantidad de zumo que contiene es inferior a cierta cantidad. Determinar cuál debe ser esa cantidad si se desea que el llenado de las botellas sean considerado como defectuoso sólo en el 5 % de los casos. (Sol: 1.835515).