eman ta zabal zazu

Universidad del País Vasco — Euskal Herriko Unibertsitatea

TAREA 1

EJERCICIOS[1]

Puedes a tu elección trabajar en R, S-PLUS o MATLAB (pero no todas las cosas que se piden pueden resultar fáciles en éste último; los ejercicios explotan ocasionalmente características de R que MATLAB no comparte).

1. Escribe una expresión que calcule Prob$\{X = x\}$ en la distribución binomial $b(p, n)$, es decir,

$$\binom{n}{x} p^x (1-p)^{n-x}$$

para valores $n = 10$, $x = 5$ y $p = 0{,}8$.

2. Escribe una función que haga lo anterior, admitiendo valores cualesquiera de n, x y p como argumentos.

3. Escribe una función que, dados n y p como argumentos, proporcione *todas* las probabilidades correspondientes a $x = 0, \ldots, n$.

4. Escribe una función bivalente que refunda las dos anteriores. Si se proporciona el argumento x, debe devolver sólo la probabilidad pedida; en otro caso, la tabla completa.

5. A partir de la identidad[2]

$$\binom{n}{x} p^x (1-p)^{n-x} = p\binom{n-1}{x-1} p^{x-1}(1-p)^{n-x+1} + (1-p)\binom{n-1}{x} p^x (1-p)^{n-x-1},$$

escribe una función como la indicada en el ejercicio 2, pero recursiva.

6. A menudo en el contraste estadístico de hipótesis, en lugar de fijar una región crítica por anticipado, se computa un valor crítico o *p-value*. Se razona así: *Supongamos* (hipótesis nula) que la variable aleatoria X se genera al muestrear una binomial $b(p, n)$, lo que denotamos por $X \sim b(p, n)$. Entonces, deberíamos observar la mayor parte de las veces un valor no muy alejado de np. Si el valor de X que observamos está “muy alejado” de np, deberemos entenderlo como evidencia contra nuestra suposición inicial (= hipótesis nula) de que $X \sim b(p, n)$.

Se define el *p-value* como la probabilidad bajo la hipótesis nula de obtener un resultado igual o más raro que el obtenido. Si la hipótesis nula fuera, por ejemplo, que la probabilidad de “cara” de una moneda es $p = p_0$ y al realizar n lanzamientos obtenemos x “caras”, el *p-value* sería:

[1] Copia de estos enunciados y, en general, de todo el material escrito repartido en clase, está disponible en `http://etdx01.bs.ehu.es`.

[2] En palabras: “La probabilidad de obtener x éxitos en n ensayos es la probabilidad de obtener $x-1$ éxitos en los primeros $n-1$ ensayos y un éxito en el n-ésimo más la probabilidad de obtener x éxitos en los primeros $n-1$ ensayos y un fracaso en el n-ésimo”.

$$\sum_{z \in \mathcal{C}} \binom{n}{z} p_0^z (1-p_0)^{n-z} \tag{1}$$

donde $\mathcal{C}$ es el conjunto de posibles resultados cuya probabilidad es menor o igual[3] que

$$\binom{n}{x} p_0^x (1-p_0)^{n-x}$$

Escribe una función que admitiendo como argumentos n, x y p_0 proporcione como resultado el *p-value* asociado a la hipótesis nula $H_0 : p = p_0$.

7. Una operación que muy frecuentemente hay que realizar es evaluar un polinomio como $p(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1}x + a_n$. ¿Cuantas operaciones (sumas, productos, cocientes) necesitas para evaluarlo de la manera más directa?

   Puede hacerse de manera más simple (regla de Horner: explicada, por ej., en [1], y en cualquier libro de Análisis Numérico). Imagina un polinomio cúbico. Podrías reescribirlo así:

$$p(z) = ((a_0 z + a_1)z + a_2)z + a_3$$

   lo que sugiere un modo de evaluarlo sin tener que calcular potencias, sólo mediante sumas y productos. Escribe una función que admita como argumentos un vector de coeficientes $a_0, \ldots, a_n$ y el valor z y proporcione $p(z)$.

AYUDAS, SUGERENCIAS Y COMPLEMENTOS

1. En R tienes funciones como `choose` y `lchoose` que permiten calcular fácilmente coeficientes binomiales y sus logaritmos (si quisieras los factoriales involucrados, podrías emplear las funciones `gamma` y `lgamma`, que computan respectivamente $\Gamma(x)$ y $\ln\Gamma(x)$; nota que $\Gamma(x) = (x-1)!$ cuando x es un número natural).

   Observa, sin embargo, que emplear las funciones anteriores reinicia siempre el cálculo *ex novo*. Las igualdades

$$\binom{n}{x} = \frac{n!}{(n-x)!x!} = \frac{n!}{(n-x+1)!(x-1)!} \times \frac{n-x+1}{x} = \binom{n}{x-1} \times \frac{n-x+1}{x},$$

   sugieren un modo fácil de computar $\binom{n}{x}$ dado $\binom{n}{x-1}$. Puedes organizar los cálculos en un bucle para aprovechar esta relación de recurrencia.

2. En realidad, la recurrencia anterior se extiende de modo obvio a *toda* la expresión $\binom{n}{x}p^x(1-p)^{n-x}$, que puede calcularse fácilmente a partir de otra precedente.

3. Otras funciones que puede interesarte mirar: `which` y `length`.

Referencias

[1] G. Dahlquist and Åke Björck. *Numerical Methods*. Prentice Hall, Englewood Cliffs, N.J., 1974.

[2] R.A. Thisted. *Elements of Statistical Computing*. Chapman & Hall, New York, 1988.

[3] Se supone un contraste sin alternativa preespecificada: si la alternativa fuera unilateral, la definición se modificaría del modo obvio.