Barcelona, a 7 de febrero de 2013

Examen de Introducción a la Probabilidad

Atención: es preciso definir previamente los sucesos relevantes. Todas las variables aleatorias utilizadas se han de especificar con claridad. Los cálculos tendrán que estar debidamente justificados.

- **Problema 1** Un entomólogo halla un ejemplar de una determinada especie de escarabajos. Dicha especie consta de tres subespecies: dos bastante abundantes (65 % y 33 % respectivamente) y una tercera, rara, constituida sólo por el 2 % de los escarabajos de la citada especie. El ejemplar encontrado luce un patrón especial de coloración de su espalda, patrón que poseen un 20 % de los individuos de la primera subespecie y sólo el 10 % de la segunda, mientras que en la subespecie rara es frecuente, presentándolo 3 de cada 4 escarabajos.

  - De cada 1000 escarabajos de la citada especie, ¿cuántos esperaríamos que tuvieran este patrón especial de coloración en su espalda? (3 pts.)

    Vamos a definir primero los acontecimientos principales:

    - E: el coleóptero presenta un patrón Especial en la espalda
    - S1: subespecie más abundante. P(S1) = 0.65. También sabemos que P(E | S1) = 0.20
    - S2: la siguiente subespecie más abundante. P(S2) = 0.33, y P(E | S2) = 0.10
    - S3: subespecie más rara. P(S3) = 1 − 0.65 − 0.33 = 0.02, y P(E | S3) = 0.75

    Se nos pregunta por P(E) o, más bien, por 1000× P(E). Por la ley de probabilidades totales:

    P(E) = P(E | S1)P(S1) + P(E | S2)P(S2) + P(E | S3)P(S3) =
    0.20*0.65+0.10*0.33+0.75*0.02 = 0.178

    Luego la respuesta es 178 escarabajos esperados (de cualquier subespecie).

  - Si una persona encuentra por azar un ejemplar de dicha especie, calcule la probabilidad de que éste sea de la subespecie más abundante y que presente además el citado patrón de coloración de su espalda. (3 pts.)

    P(S1 ∩ E) = P(E | S1)P(S1) = 0.20*0.65 = 0.13

    Las dos características coinciden en un 13 % de los ejemplares.

  - Teniendo en cuenta que el ejemplar que encontró el entomólogo presentaba el patrón antes mencionado, ¿cuál es la probabilidad de que pertenezca a la subespecie rara? (4 pts.)

    Formulamos la pregunta como probabilidad condicionada, ya que hemos de limitarnos solo al conjunto de escarabajos que presentan el patrón.

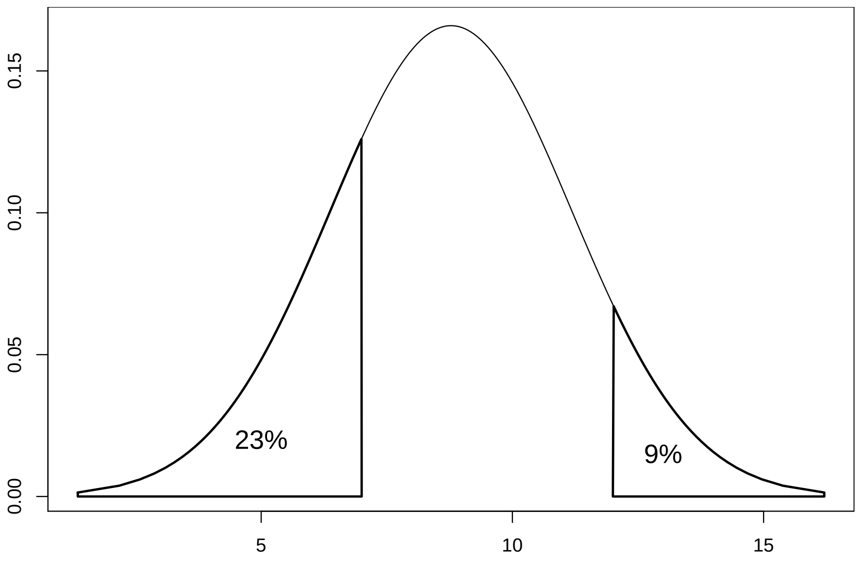
    $$P(S3|E) = \frac{P(S3 \cap E)}{P(E)} = \frac{0.015}{0.178} = 0.08427$$

    El hecho de tener el patrón ha incrementado la probabilidad de ser de la tercera subespecie del 2 % al 8.4 %.

- **Problema 2** Sabemos que la masa de los individuos machos de una determinada especie de coleópteros sigue, con gran aproximación, una distribución Normal; que el 23 % de dichos individuos posee una masa menor de 7 gramos, y para el 9 % de los mismos su masa excede de 12 gramos.

  - Represente en la figura, con la mayor precisión posible, la información precedente. (2 pts.)
  - Deduzca cuáles son los valores de la esperanza y la desviación estándar de la masa de los citados escarabajos (deducirlo de la gráfica no sirve, es preciso encontrarlo a partir de de los datos). (2 pts.)

    Planteamos la información disponible: si $M \sim \mathrm{N}(\mu, \sigma)$, es la masa del coleóptero,



- P(M <7) = 0.23, es decir, $\mu + Z_{0.23}\sigma = 7$
- P(M >12) = 0.09, es decir, $\mu + Z_{0.91}\sigma = 12$

De las tablas encontramos que el percentil 0.91 de una N(0,1) vale aproximadamente 1.34. De hecho, lo que podemos encontrar generalmente es que para el valor 1.34 su probabilidad es 0.9099, que está muy cerca de la probabilidad que tenemos. Para el percentil 0.23 se ha de considerar que será un valor negativo, y que su opuesto es el percentil $1 - 0.23 = 0.77$, aproximadamente 0.74 (probabilidad exacta para 0.74: 0.77035). Por tanto, planteamos un sistema de dos ecuaciones, donde las incógnitas son μ y σ:

$$\mu - 0.74\sigma = 7$$
$$\mu + 1.34\sigma = 12$$

La solución es $\mu = 8.779$ y $\sigma = 2.404$. Comprobamos que, de acuerdo a la figura, tienen pleno sentido.

- Calcule la probabilidad de que un escarabajo macho de la citada especie hallado al azar tenga una masa comprendida entre 8 y 14 gramos. Utilice los parámetros calculados en el apartado anterior, si tienen sentido. En caso contrario, asuma 9 y 2 gramos para la esperanza y la desviación estándar respectivamente. (3 pts.)

  P(8 <M <14) = P($\frac{8-\mu}{\sigma} < Z < \frac{14-\mu}{\sigma}$) = P(-0.324 <Z <2.172)
  El valor de $F_Z(2.172)$ (usando 2.17) es aproximadamente 0.985, y el valor de $F_Z(-0.324) = 1 - F_Z(0.324)$ lo estimaremos a partir de los valores en la tabla de 0.32 y 0.33: respectivamente, 0.6255 y 0.6293. Ponderando convenientemente obtenemos que $F_Z(-0.324) = 1 - 0.6270 = 0.373$. Luego la probabilidad pedida es 0.985-0.373=0.612.

- Una trampa ha capturado 12 escarabajos macho de la citada especie durante la pasada noche. ¿Cuál es la probabilidad de que el peso total de estos 12 animales sea inferior a 90 gramos? (3 pts.)

  Si pensamos en grupos de 12 escarabajos de este tipo y su peso total T, tenemos una distribución Normal con:

  - media $12\mu = 105.348$ gramos
  - variancia $12\sigma^2 = 69.34$ o, lo que es lo mismo, $\sigma = 8.327$ gramos.

  Luego la pregunta planteada es P(T <90):

$$P(T < 90) = P\left(Z < \frac{90 - 105.348}{8.327}\right) = F_Z(-1.842) = 1 - F_Z(1.842) \approx 0.0329$$

- **Problema 3** Sea X una variable aleatoria absolutamente continua con función de densidad de probabilidad

$$f(x) = \begin{cases} \alpha x^3 & \text{si } 0 \le x \le 2 \\ 0 & \text{en caso contrario} \end{cases}$$

Mostrando explícitamente los cálculos:

  - Hallar el valor de α (2.5 pts.)

    Puesto que f es una función de densidad, tendremos:

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{2} f(x)\,dx + \int_{2}^{\infty} f(x)\,dx \\ &= \int_{-\infty}^{0} 0\,dx + \int_{0}^{2} \alpha x^3\,dx + \int_{2}^{\infty} 0\,dx = 0 + \alpha\left[\frac{x^4}{4}\right]_0^2 + 0 = \alpha\,\frac{2^4}{4} = 4\alpha \Longrightarrow \alpha = \frac{1}{4} \end{aligned}$$

  - Hallar la función de distribución de X. (2.5 pts.)

    Si llamamos F_X a la función de distribución de X, entonces, al tomar X valores entre 0 y 2 con probabilidad 1, resulta que para $x \le 0$ tendremos:

$$F_X(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{x} 0\,dt = 0$$

para $x \in [0,2]$ tendremos:

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{0} f(t)\,dt + \int_{0}^{x} f(t)\,dt = \int_{-\infty}^{0} 0\,dt + \int_{0}^{x} \frac{1}{4}t^3\,dt \\ &= 0 + \frac{1}{4}\left[\frac{t^4}{4}\right]_0^x = \frac{1}{16}x^4 \end{aligned}$$

para $x > 2$ tendremos:

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{0} f(t)\,dt + \int_{0}^{2} f(t)\,dt + \int_{2}^{x} f(t)\,dt \\ &= \int_{-\infty}^{0} 0\,dt + \int_{0}^{2} \frac{1}{4}t^3\,dt + \int_{2}^{x} 0\,dt = 0 + \frac{1}{4}\left[\frac{t^4}{4}\right]_0^2 + 0 = 1 \end{aligned}$$

  - Hallar la probabilidad de que $X \ge 1$. (2.5 pts.)

    Al ser X absolutamente continua, la probabilidad de un punto es cero y podemos escribir:

$$P(X \ge 1) = P(X > 1) = 1 - P(X \le 1) = 1 - F_X(1) = 1 - \frac{1}{16} = \frac{15}{16} = 0,9375$$

  - Hallar $E(X)$, $E(X^2)$ y $\mathrm{var}(X)$. (2.5 pts.)

    Tendremos:

$$E(X) = \int_{-\infty}^{\infty} x f(x)\,dx = \int_{0}^{2} \frac{1}{4}x^4\,dx = \frac{1}{4}\left[\frac{x^5}{5}\right]_0^2 = \frac{2^5}{20} = \frac{32}{20} = \frac{8}{5} = 1,6$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_{0}^{2} \frac{1}{4}x^5\,dx = \frac{1}{4}\left[\frac{x^6}{6}\right]_0^2 = \frac{2^6}{24} = \frac{8}{3} = 2,6667$$

$$\mathrm{var}(X) = E(X^2) - E(X)^2 = \frac{8}{3} - \left(\frac{8}{5}\right)^2 = \frac{200-192}{75} = \frac{8}{75} = 0,1067$$

- **Problema 4** En una población hay 10000 jóvenes, se estima que la probabilidad de que un joven vaya al cine un determinado fin de semana es igual a 0,1. Sea X la variable aleatoria *número de jóvenes que han ido al cine el pasado fin de semana*.

  - ¿Cuál es la distribución de X? ¿Cuales son la esperanza y la varianza de X? (2 pts.)
    En las condiciones del enunciado y asumiendo independencia de los sucesos ir al cine de cada uno de los jóvenes, X seguirá una distribución Binomial, $B(n = 10000, p = 0,1)$. Su esperanza será $E(X) = np = 10000 \times 0,1 = 1000$ y su varianza $\text{var}(X) = np(1-p) = 10000 \times 0,1 \times (1-0,1) = 900$.
  - Si elegimos 5 jóvenes de la población en condiciones de independencia ¿cuál es la probabilidad exacta de que sólo haya ido al cine como máximo uno de ellos? (2 pts.)
    Sea Y la variable aleatoria número de jóvenes entre estos cinco que ha ido al cine. En las condiciones del enunciado, Y seguirá una distribución Binomial, $B(n = 5, p = 0,1)$. Nos piden $P(Y \leq 1)$. Teniendo en cuenta que

$$P(Y = k) = \binom{5}{k} 0,1^k 0,9^{(5-k)}$$

    resulta que

$$P(Y \leq 1) = P(Y = 0) + P(Y = 1) = 0,9^5 + 5 \times 0,1 \times 0,9^4 = 0,91854$$

    .
  - Halle la probabilidad de $920 \leq X \leq 1040$, de forma aproximada. (3 pts.)
    Usaremos el Teorema de Laplace-Moivre (o el Teorema del Límite Central) para aproximar la distribución de X (una Binomial) a una Normal, puesto que $n \geq 30$ con $np \geq 3$ y $n(1-p) \geq 3$. En otras palabras consideraremos que la variable aleatoria

$$Z = \frac{X - np}{\sqrt{np(1-p)}} = \frac{X - 1000}{30}$$

    sigue aproximadamente una distribución normal, N(0,1). Por tanto

$$\begin{aligned} P(920 \leq X \leq 1040) &= P\left(\tfrac{920-1000}{30} \leq \tfrac{X-1000}{30} \leq \tfrac{1040-1000}{30}\right) = P\left(-\tfrac{8}{3} \leq Z \leq \tfrac{4}{3}\right) \\ &= F_Z\left(\tfrac{4}{3}\right) - F_Z\left(-\tfrac{8}{3}\right) = F_Z\left(\tfrac{4}{3}\right) - \left(1 - F_Z\left(\tfrac{8}{3}\right)\right) = F_Z\left(\tfrac{4}{3}\right) + F_Z\left(\tfrac{8}{3}\right) - 1 \\ &= F_Z\big(1,3333\big) + F_Z\big(2,6667\big) - 1 \approx 0,9088 + 0,9962 - 1 = 0,9050 \end{aligned}$$

  - Supongamos que escogemos jóvenes al azar de dicha población, en condiciones de independencia estocástica, hasta encontrar a uno que haya ido al cine el pasado fin de semana. ¿Cuál es la probabilidad de que hayamos tenido que escoger a cinco jóvenes? ¿Cuál es el número medio de jóvenes escogidos? (3 pts.)
    Sea U la variable aleatoria número total de jóvenes escogidos que no han ido al cine hasta encontrar uno que haya ido. En las condiciones del enunciado U seguirá una distribución de Pascal o Geométrica de parámetro p. Por tanto, la probabilidad buscada es:

$$P(U + 1 = 5) = P(U = 4) = 0,9^4 \times 0,1 = 0,06561$$

    Además, el valor medio buscado es:

$$E(U + 1) = E(U) + 1 = \frac{1-p}{p} + 1 = \frac{1}{p} = \frac{1}{0,1} = 10$$