Clase 11 de Bioestadística

Esperanza de variables aleatorias continuas

Ernesto Mordecki

CMAT, Facultad de Ciencias, Universidad de la República.

Uruguay

Contenidos de la clase

- Esperanza (repaso)
- Variable uniforme
  - Aplicación: Ley fuerte de los grandes números (LFGN)
- Esperanza de una variable aleatoria normal
- Esperanza de una variable exponencial
- Aplicación: la medusa inmortal

Esperanza (repaso)

Definición: Llamamos esperanza matemática de una variable aleatoria X al número $E(X)$ definido como

$$E(X) = \sum_k x_k P(X = x_k)$$

si X es una variable discreta y

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

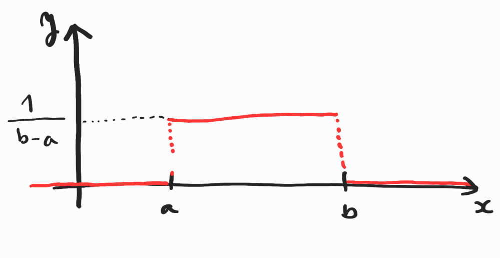
donde f es la densidad de X

Para calcular la esperanza matemática de una variable continua tenemos que conocer la la densidad de la variable aleatoria X.

Ejemplo: variable unforme en $[a, b]$

- Si U es una variable uniforme en $[a, b]$ su densidad vale

$$f(x) = \frac{1}{b-a}, \qquad a \leq x \leq b.$$



- ¿dónde debería estar ubicada la esperanza?
- ¿cuánto debería valer?

Hagamos las cuentas:

Partimos de la definición:

$$E(U) = \int_{-\infty}^{\infty} x\, f(x) dx = \int_a^b x\, f(x) dx$$

$$= \int_a^b x\, \frac{1}{b-a} dx = \frac{1}{b-a} \int_a^b x\, dx$$

$$= \frac{1}{b-a} \left[\frac{x^2}{2}\right]_{x=a}^{x=b} = \frac{1}{b-a} \left[\frac{b^2-a^2}{2}\right]$$

$$= \frac{1}{\cancel{b-a}} \left[\frac{\cancel{(b-a)}(b+a)}{2}\right] = \frac{a+b}{2}$$

Aplicación: Ley fuerte de los grandes números (LFGN)

- Sea $U_1, U_2, \ldots$ una sucesión de variables aleatorias independientes e idénticamente distribuídas, uniformes en el intervalo $[1, 3]$
- Nos intersa aproximar el promedio

$$\overline{U}_n = \frac{U_1 + \cdots + U_n}{n}$$

- Aplicamos la LFGN: sabemos que $E(U_1) = 2$.
- Obtenemos

$$\overline{U}_n \sim 2$$

- La calidad de la aproximación dependerá de n (y del azar)

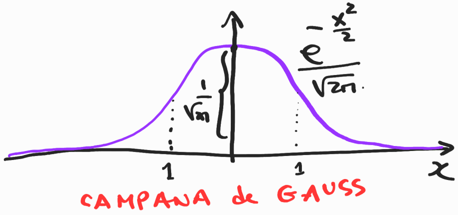
Simulamos en R

```
>
> mean(runif(10,min=1,max=3))
[1] 1.8904
> mean(runif(100,min=1,max=3))
[1] 2.00482
> mean(runif(1000,min=1,max=3))
[1] 1.993221
> mean(runif(10000,min=1,max=3))
[1] 2.002744
> mean(runif(100000,min=1,max=3))
[1] 2.002927
> mean(runif(1e6,min=1,max=3))
[1] 1.999075
> mean(runif(1e7,min=1,max=3))
[1] 2.000061
> |
```

Esperanza de una variable aleatoria normal

- Comenzamos con una variable normal estándar Z, es decir $Z \sim \mathcal{N}(0,1)$ y su densidad vale[1]

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \qquad x \in \mathbb{R}.$$



- ¿dónde debería estar ubicada la esperanza?
- ¿cuánto debería valer?

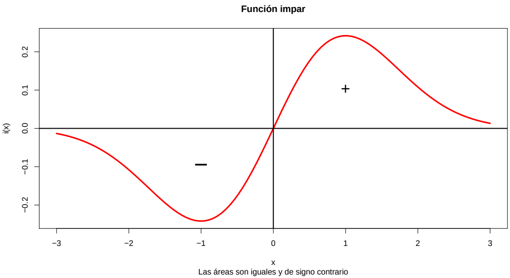
[1] Se utiliza φ para la densidad de $Z \sim \mathcal{N}(0,1)$

Vamos a calcular:

$$E(Z) = \int_{-\infty}^{\infty} x\ \varphi(x) dx = \int_{-\infty}^{\infty} x\ \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x\ e^{-x^2/2} dx = 0,$$

por simetría (el integrando es una función impar):



El código que me produce el gráfico

```
# grafico una función
# lo primero es definirla
i<-function(x) {x*dnorm(x)}
# le indico donde quiero guardar el pdf que produzco
setwd("/home/mordecki/Dropbox/1021_Bioestadistica
      /clases_teoricas/figuras")
# este comando abre un PDF de tamaño 11 x 6
pdf("impar.pdf",width = 11, height = 6)
# el comando "curve" grafica una función
curve(i,-3,3,lwd=3,col="red",main="Función impar",
      sub="Las áreas son iguales y de signo contrario")
# así dibujo los ejes:
abline(h=0,lwd=2)
abline(v=0,lwd=2)
# ahora agrego texto en el dibujo
text(1,0.1,"+",cex=2)
text(-1,-0.1,"-",cex=3)
# por último cierro el PDF y queda guardado en el lugar indicado
dev.off()
```

Esperanza de una normal (μ, σ^2)

Para pasar al caso general, recordamos el resultado de estandarización

$$X \sim \mathcal{N}(\mu, \sigma^2) \quad \text{si y solo si} \quad Z = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

¿Como se traduce este resultado en términos de densidades?

- Sea $\varphi(x)$ la densidad de Z y $f(x)$ la densidad de X. Tenemos

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \qquad f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)\right]^2$$

- Es decir

$$f(x) = \frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right)$$

- Donde decía z pusimos $\frac{x-\mu}{\sigma}$. Es un cambio de variable.
- Ahora podemos calcular $E(X)$ para $X \sim \mathcal{N}(\mu, \sigma^2)$

- Según la definición

$$E(X) = \int_{-\infty}^{\infty} x\, f(x)dx = \int_{-\infty}^{\infty} x\, \frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right) dx$$

- Ahora cambiamos la variable en la integral de acuerdo a

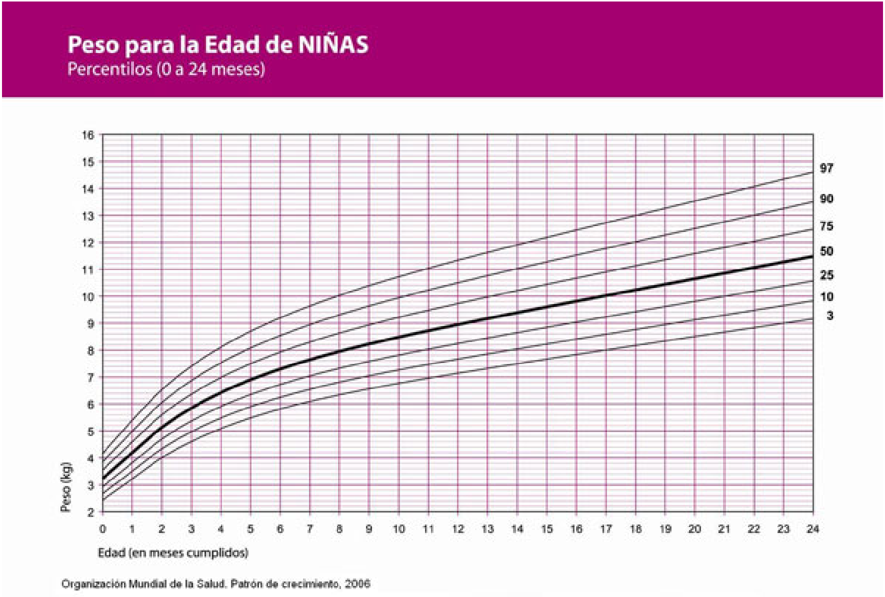
$$z = \frac{x-\mu}{\sigma}, \quad dz = \frac{1}{\sigma}dx, \quad x = \sigma z + \mu.$$

- Los límites de integración no cambian. Entonces

$$E(X) = \int_{-\infty}^{\infty} (\sigma z + \mu)\, \varphi(z)dz$$

$$= \sigma \boxed{\int_{-\infty}^{\infty} z\, \varphi(z)dz} + \mu \boxed{\int_{-\infty}^{\infty} \varphi(z)dz} = \mu.$$

Comentario: pesos de las bebés



Esperanza de una variable exponencial

- En este caso T tiene densidad

$$f(t) = \lambda e^{-\lambda t}$$

- Entonces

$$E(T) = \int_{-\infty}^{\infty} t\, f(t) dt = \lambda \int_{0}^{\infty} t\, e^{-\lambda t} dt$$

- Cambiamos de variable

$$t = \frac{v}{\lambda}, \quad dt = \frac{dv}{\lambda}, \quad v = \lambda t,$$

- Obtenemos

$$E(T) = \lambda \int_{0}^{\infty} \frac{v}{\lambda}\, e^{-v} \frac{dv}{\lambda} = \frac{1}{\lambda} \boxed{\int_{0}^{\infty} v\, e^{-v} dv}$$

La integral en la caja se calcula por partes:

$$\int_a^b f(v)g'(v)dv = f(b)g(b) - f(a)g(a) - \int_a^b f'(v)g(v)dv$$

Tomamos entonces

$$f(v) = v, \quad g'(v) = e^{-v}$$

$$f'(v) = 1, \quad g(v) = -e^{-v}$$

Obtenemos

$$\int_0^\infty v\ e^{-v}dv = \lim_{v\to\infty}(-ve^{-v}) - 0e^{-0} - \int_0^\infty(-e^{-v})dv$$

$$= \int_0^\infty e^{-v}dv = \left[-e^{-v}\right]_{v=0}^{v=\infty} = 1.$$

Conclusión

$$\boxed{E(T) = \frac{1}{\lambda}.}$$

Aplicación: la medusa inmortal

- La medusa inmortal (Turritopsis Nutricula) se considera bioógicamente inmortal.
- Dado que se reproducen, de ser inmortales su cantidad aumentaría exponencialmente.
- Consideramos que mueren entonces por predación o por enfermedades.
- Debido a que no envejecen, podemos suponer que tienen un tiempo de vida exponencial
- Suponiendo que en promedio viven un siglo, ¿cuál es la probabilidad de que un ejemplar viva mas de 500 años?

Solución[2]

- Si T representa el tiempo de vida de un ejemplar, tenemos, medido en años:

$$E(T) = \frac{1}{\lambda} = 100.$$

- Tenemos que calcular

$$P(T > 500) = e^{-\lambda \times 500} = e^{-500/100} = e^{-5} = 0{,}007$$

[2]Ojo: el ejemplo es una especulación: