

Modelos y simulación - Ejercicios resueltos

FAMAF - UNC

Severino Di Giovanni



Figure 1: Severino Di Giovanni, el autor de este apunte. Un anarquista libertario, murió luchando por la libertad. Como él, otros miles han muerto para que nosotros gocemos de los derechos que tenemos. No te dejes engañar por los tristes pregoneros del egoísmo. Amá a tu prójimo y no olvides que si sus derechos se vulneran, los tuyos también. Ayudá a tu compañero de estudio, defendé tu universidad.

Contents

1	P1	3
2	P3	10

1 P1

(1) Un geólogo recoge 10 especímenes de A y 12 de B. 15 son elegidos al azar. De la función de probabilidad de masa del número de especímenes de tipo A elegidos.

Sea Ω el espacio muestral, con cada $\omega \in \Omega$ una de las posibles selecciones de 15 de entre los 22 especímenes. Claramente,

$$|\Omega| = \binom{22}{15}$$

Sea X la variable aleatoria para el número de especímenes de tipo A en una selección aleatoria de Ω . Claramente, para tener $X = k$, con $3 \leq k \leq 10$, la selección debe ser tal que k muestras sean de tipo A y $15 - k$ muestras sean de tipo B. Luego

$$P(X = k) = \frac{\binom{10}{k} \binom{12}{15-k}}{\binom{22}{15}}$$

(2) Sean X, Y independientes y exponenciales con parámetros λ, μ respectivamente. Computar $f_{X|Y}(x | y)$ y $P(X < Y)$.

(a) Recordemos que $f_{X|Y}(x | y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$. Como X, Y son independientes, su densidad conjunta es el producto de las densidades, y tenemos $f_{X|Y}(x | y) = (f_X(x)f_Y(y))/f_Y(y) = f_X(x)$.

(b) Veamos que

$$\begin{aligned} P(X < Y) &= \int_{\mathbb{R}} f_Y(y) \int_{-\infty}^y f_X(x) dx dy \\ &= \int_{\mathbb{R}} \mathcal{I}_{(0,\infty)} \cdot \mu e^{-\mu y} \int_{-\infty}^y \mathcal{I}_{(0,\infty)} \lambda e^{-\lambda x} dx dy \\ &= \int_0^{\infty} \mu e^{-\mu y} \int_0^y \lambda e^{-\lambda x} dx dy \end{aligned}$$

Ahora bien, con $w := -\lambda x, dw = -\lambda dx$, tenemos

$$\begin{aligned} \int_0^y \lambda e^{-\lambda x} dx &= - \int_0^{w(y)} e^w dw \\ &= - \int_0^{-\lambda y} e^w dw \\ &= - [e^{-\lambda y} - e^0] \\ &= 1 - e^{-\lambda y} \end{aligned}$$

Por lo tanto

$$\begin{aligned} P(X < Y) &= \int_0^{\infty} \mu e^{-\mu y} (1 - e^{-\lambda y}) dy \\ &= \int_0^{\infty} \mu e^{-\mu y} dy - \int_0^{\infty} \mu e^{-(\mu+\lambda)y} dy \\ &= \int_0^{\infty} f_Y(y) dy - \frac{\mu}{\mu+\lambda} \int_0^{\infty} (\mu+\lambda) e^{-(\mu+\lambda)y} dy \\ &= 1 - \frac{\mu}{\mu+\lambda} \int_0^{\infty} f_Z(y) dy \quad \{Z \sim \mathcal{E}(\mu+\lambda)\} \\ &= 1 - \frac{\mu}{\mu+\lambda} \\ &= \frac{\mu+\lambda}{\mu+\lambda} - \frac{\mu}{\mu+\lambda} \\ &= \frac{\lambda}{\mu+\lambda} \end{aligned}$$

- (3) Sean X, Y independientes y exponenciales con el mismo parámetro λ . Computar la densidad condicional de X dado que $X + Y = t$.

Por def. de densidad condicional, tomando $T := X + Y$,

$$f_{X|T}(x | t) = \frac{f_{X,T}(x, t)}{f_T(t)}$$

Calculemos estas distribuciones.

$$\begin{aligned} f_{X,T}(x, t) &= P(X = x \cap X + Y = t) \\ &= P(X = x \cap Y = t - x) \\ &= P(X = x) \cdot P(Y = t - x) && \{X, Y \text{ son independientes}\} \\ &= f_X(x) \cdot f_Y(t - x) \\ &= \lambda e^{-\lambda x} \mathcal{I}_{(0, \infty)}(x) \cdot \lambda e^{-\lambda(t-x)} \mathcal{I}_{(0, \infty)}(t - x) \\ &= \lambda^2 e^{-\lambda t} \mathcal{I}_{(0, \infty)}(t - x) \mathcal{I}_{(0, \infty)}(x) \\ &= \lambda^2 e^{-\lambda t} \mathcal{I}_{(x, \infty)}(t) \mathcal{I}_{(0, \infty)}(x) \end{aligned}$$

El último paso se justifica porque $\mathcal{I}_{(0, \infty)}(t-x) \mathcal{I}_{(0, \infty)}(x) = \mathcal{I}_{(0, \infty)}(x) \mathcal{I}_{(x, \infty)}(t)$. En efecto, $t - x$ pertenece a $(0, \infty)$ si y solo si $t > x$. Ahora bien,

$$\begin{aligned} f_T(t) &= f_{X+Y}(t) \\ &= \int_{-\infty}^t f_X(x) f_Y(t - x) dx && \{X, Y \text{ independientes}\} \\ &= \int_{-\infty}^t \mathcal{I}_{(0, \infty)}(x) \cdot \lambda e^{-\lambda x} \cdot \mathcal{I}_{(0, \infty)}(t - x) \lambda e^{-\lambda(t-x)} dx \\ &= \int_0^t \lambda^2 e^{-\lambda t} \cdot \mathcal{I}_{(x, \infty)}(t) dx \\ &= \lambda^2 e^{-\lambda t} \int_0^t \mathcal{I}_{(x, \infty)}(t) dx \\ &= \lambda^2 e^{-\lambda t} \int_0^t 1 dx && (\star) \\ &= t \lambda^2 e^{-\lambda t} \end{aligned}$$

El paso a (\star) se justifica como sigue. Si la región de integración es $(0, t)$ y la variable de integración es x , siempre se cumple que $x < t$, y por lo tanto $\mathcal{I}_{(x, \infty)}(t) = 1$.

$$\therefore f_{X|T}(x | t) = \frac{f_{X,T}(x, t)}{f_T(t)} = \frac{\lambda^2 e^{-\lambda t} \mathcal{I}_{(0, \infty)(x)} \mathcal{I}_{(x, \infty)}(t)}{t \lambda^2 e^{-\lambda t}} = \begin{cases} \frac{1}{t} & 0 \leq x \leq t \\ 0 & c.c. \end{cases}$$

(4) Sean $N_1(t), N_2(t)$ procesos con $t \geq 0$ y tasas $\lambda = 1, \mu = 2$ respectivamente. Sea $N(t) = N_1(t) + N_2(t)$.

- (a) Calcular la probabilidad de que $N(1) = 2, N(2) = 5$.
- (b) Calcular la probabilidad de que $N_1(1) = 1$ dado $N(1) = 2$.
- (c) Calcular la probabilidad de que el segundo arribo en N_1 ocurra antes que el tercer arribo en $N_2(t)$.

(a) Es muy fácil porque $N(t)$ es un proceso de Poisson con parámetro $\lambda + \mu = 3$, con lo cual $P(N(t) = k) = e^{-3} \frac{3^k}{k!}$ para cualquier $k \geq 0$ entero.

(b) Claramente, si $N(1) = 2$, la probabilidad de que $N_1(1) = 1$ es la probabilidad de que $N_2(1) = 1$.

$$\begin{aligned}
 P(N_1(1) = 1 \mid N(1) = 2) &= \frac{P(N_1(1) = 1 \cap N(1) = 2)}{P(N(1) = 2)} \\
 &= \frac{P(N_1(1) = 1 \cap N_1(1) + N_2(1) = 2)}{P(N(1) = 2)} \\
 &= \frac{P(N_1(1) = 1 \cap N_2(1) = 1)}{P(N(1) = 2)} \\
 &= \frac{P(N_1(1) = 1) \cdot P(N_2(1) = 1)}{P(N(1) = 2)} \quad \{\text{Independencia de } N_1, N_2\}
 \end{aligned}$$

Ahora solo resta ver que

$$e^{-\lambda} \frac{(\lambda l)^k}{k!} \leq e^{-\mu} \frac{(\mu l)^k}{k!} \iff e^{-\lambda l} \lambda^k \leq e^{-\lambda \mu} \mu^k \iff \lambda \ln \lambda \geq \mu \ln \mu$$

$$P(N_1(1) = 1) = e^{-1} \frac{1^1}{1!} = e^{-1} \approx 0.368$$

$$P(N_2(1) = 1) = e^{-2} \frac{2^1}{1!} = 2e^{-2} \approx 0.27$$

$$P(N(1) = 2) = e^{-3} \frac{3^2}{2!} = \frac{9}{2} e^{-3} \approx 0.224$$

$$\therefore P(N_1(1) = 1 \mid N(1) = 2) \approx \frac{0.368 \cdot 0.27}{0.224} = 0.443$$

Otra forma de calcularlo es usando el teorema de Bayes:

$$\begin{aligned}
P(N_1(1) = 1 \mid N(1) = 2) &= \frac{P(N(1) = 1 \mid N_1(1) = 1)P(N_1(1) = 1)}{P(N(1) = 1)} \\
&= \frac{P(N_1(1) + N_2(1) = 1 \mid N_1(1) = 1)P(N_1(1) = 1)}{P(N(1) = 1)} \\
&= \frac{P(N_2(1) = 1) \cdot P(N_1(1) = 1)}{P(N(1) = 2)}
\end{aligned}$$

que es lo mismo que tuvimos antes.

(c) Se nos pide la probabilidad de que el segundo arribo en $N_1(t)$ ocurra antes que el tercer arribo en $N_2(t)$. A primera vista no parece, pero este es simplemente un problema de conteo.

Primero, observemos que el evento "el segundo arribo en $N_1(t)$ ocurre antes que el tercero de $N_2(t)$ " es equivalente al evento "de 4 eventos en $N(t)$, al menos dos provienen de $N_1(t)$ ". Sea ω este evento.

La probabilidad de que un evento dado de $N(t)$ provenga de $N_1(t)$ es $\frac{\lambda}{\lambda+\mu}$. La probabilidad de que un evento dado de $N(t)$ provenga de $N_2(t)$ es $\frac{\mu}{\lambda+\mu}$.

$$\therefore P(\omega) = \binom{4}{2} \frac{\lambda^2 \mu^2}{(\lambda + \mu)^4} + \binom{4}{3} \frac{\lambda^3 \mu}{(\lambda + \mu)^4} + \binom{4}{4} \frac{\lambda^4}{(\lambda + \mu)^4}$$

Sustituyendo λ, μ por sus valores y calculando los binomiales, que son fáciles, sale el resultado.

(5) Sean X, Y independientes con distribuciones de Poisson parametrizadas bajo λ, μ respectivamente. Probar que $Z = X + Y \sim \mathcal{P}(\lambda + \mu)$.

Como son independientes, la distribución de su suma es dada por la "convolución" discreta de sus distribuciones. Si con \mathbb{N} denotamos los naturales y el cero, entonces

$$\begin{aligned}
 f_Z(z) &= \sum_{k \in \mathbb{N}} p_X(k) p_Y(z-k) \\
 &= \sum_{k \in \mathbb{N}} \left(e^{-\lambda} \frac{\lambda^k}{k!} \right) \left(e^{-\mu} \frac{\mu^{z-k}}{(z-k)!} \right) \\
 &= e^{-\lambda} e^{-\mu} \sum_{k \in \mathbb{N}} \frac{\lambda^k \mu^{z-k}}{k! (z-k)!} \\
 &= \frac{\eta}{z!} \sum_{k \in \mathbb{N}} \lambda^k \mu^{z-k} \frac{z!}{k! (z-k)!} \quad \left\{ \eta := e^{-\lambda} e^{-\mu} \right\} \\
 &= \frac{\eta}{z!} \sum_{k \in \mathbb{N}} \lambda^k \mu^{z-k} \binom{z}{k}
 \end{aligned}$$

Ahora bien, como $Y \sim \mathcal{P}(\mu)$, $\text{Im}(Y) = \mathbb{N}$, y por lo tanto $p_Y(u) = 0$ si $u < 0$. En otras palabras, dado un k arbitrario, $p_Y(z-k) \neq 0 \iff z-k \geq 0 \iff k \leq z$.

$$\therefore \frac{\eta}{z!} \sum_{k \in \mathbb{N}} \lambda^k \mu^{z-k} \binom{z}{k} = \frac{\eta}{z!} \sum_{k=0}^z \lambda^k \mu^{z-k} \binom{z}{k} = \frac{\eta}{z!} (\lambda + \mu)^z = e^{-(\lambda + \mu)} \frac{(\lambda + \mu)^z}{z!}$$

por el teorema del binomio de Newton.

$\therefore Z \sim \mathcal{P}(\lambda + \mu)$. ■

2 P3

(2) (Todas las variables en este problema son uniformes continuas e independientes en $(0, 1)$). En un juego, se simula la variable aleatoria U . Si $U < \frac{1}{2}$ se suman dos números aleatorios. Si $U \geq \frac{1}{2}$, se suman tres. El resultado X de cualquiera de estas sumas es el puntaje, y se tiene éxito si $X \geq 1$.

(a) Dar la probabilidad de ganar.

Por ley de probabilidad total,

$$\begin{aligned} & P(X_1 + X_2 \geq 1 \mid U < \frac{1}{2})P(U < \frac{1}{2}) + P(X_1 + X_2 + X_3 \geq 1 \mid U \geq \frac{1}{2})P(U \geq \frac{1}{2}) \\ &= P(X_1 + X_2 \geq 1)P(U < \frac{1}{2}) + P(X_1 + X_2 + X_3 \geq 1)P(U \geq \frac{1}{2}) \\ &= P(X_1 + X_2 \geq 1) \cdot \frac{1}{2} + P(X_1 + X_2 + X_3 \geq 1) \cdot \frac{1}{2} \end{aligned}$$

Estudiemos $P(X_1 + X_2 = z)$. Veamos que si $z \in [0, 1]$, entonces

$$P(X_1 + X_2 = z) = \int_0^z f_{X_1}(x)f_{X_2}(z-x) dx = \int_0^z 1 dx = z$$

Si $z \in (1, 2]$, entonces hacemos variar X_1 entre $[z-1, 1]$ y requerimos $X_2 = z - X_1$:

$$P(X_1 + X_2 = z) = \int_{z-1}^1 f_{X_1}(x)f_{X_2}(z-x) dx = 2 - z$$

$$\therefore f_{X_1+X_2}(x) = \begin{cases} x & 0 \leq x \leq 1 \\ 2-x & 1 < x \leq 2 \\ 0 & c.c. \end{cases}$$

Por lo tanto,

$$P(X_1 + X_2 \geq 1) = \int_1^2 f_{X_1+X_2}(x) dx = \int_1^2 (2-x) dx = 1/2$$

Resta ver que

$$\begin{aligned}
P(X_1 + X_2 + X_3 \geq 1) &= P(X_1 + X_2 \geq 1 - X_3) \\
&= \int_0^1 P(X_1 + X_2 \geq 1 - z) P(X_3 = z) \, dz \quad \{\text{Por prob. total}\} \\
&= \int_0^1 \left(1 - \frac{(1-z)^2}{2}\right) \cdot 1 \, dz \quad \{\text{Por } \star\} \\
&= 5/6
\end{aligned}$$

dado que para $z \in [0, 1]$

$$(\star) \quad P(X_1 + X_2 \geq 1 - z) = 1 - P(X_1 + X_1 \leq 1 - z) = 1 - \int_0^{1-z} x \, dx = 1 - \frac{(1-z)^2}{2}$$

$$\begin{aligned}
\therefore P(X_1 + X_2 \geq 1) &\cdot \frac{1}{2} + P(X_1 + X_2 + X_3 \geq 1) \cdot \frac{1}{2} \\
&= \frac{1}{2} \cdot \frac{1}{2} + \frac{5}{6} \cdot \frac{1}{2} \\
&= \frac{1}{4} + \frac{5}{12} \\
&= \frac{2}{3}
\end{aligned}$$

(b) Implement an algorithm that estimates the probability of winning, i.e. the proportion of wins in n simulations of the game.

In Python,

```
[python]
from random import uniform

def sim_2(n):

    wins = 0

    # S = summation of uniform rvs
    def S(k):
        return sum( [uniform(0, 1) for _ in range(k)] )

    for _ in range(n):

        u = uniform(0, 1)

        X = S(2) if u < 0.5 else S(3)
        wins = wins + 1 if X >= 1 else wins

    return wins/n
```

(3) Las máquinas tragamonedas usualmente generan un premio cuando hay un acierto. Supongamos que se genera el acierto con el siguiente esquema: se genera un número aleatorio, y

- si es menor a un tercio, se suman dos nuevos números aleatorios
- si es mayor o igual a un tercio, se suman tres números aleatorios .

Si el resultado de la suma es menor o igual a 2, se genera un acierto.

a) ¿Cuál es la probabilidad de acertar?.

Sea $Z_n = \sum_{i=1}^n U_i$ con $U_i \sim \mathcal{U}(0,1)$ para $1 \leq i \leq n$. Por ley de probabilidad total, se llega con facilidad a

$$P(\text{Ganar}) = P(Z_2 \leq 2) \cdot \frac{1}{3} + P(Z_3 \leq 2) \cdot \frac{2}{3}$$

Claramente, $P(Z_2 \leq 2) = 1$, con lo cual solo queda computar $P(Z_3 \leq 2) = P(Z_2 \leq 2 - U_3)$. La densidad de Z_2 fue calculada en el punto (2). Para calcular $P(Z_3 \leq 2)$, podemos hacer variar a U_3 en todo el rango $[0, 1]$ y a Z_2 en todo el rango $[0, 2 - U_3]$:

$$\begin{aligned} P(Z_3 \leq Z_2) &= P(Z_2 \leq 2 - U_3) \\ &= \int_0^1 P(U_3 = z) \int_0^{2-z} f_{Z_2}(x) dx dz \\ &= \int_0^1 \left[\int_0^1 f_{Z_2}(x) dx + \int_1^{2-z} f_{Z_2}(x) dx \right] dz \\ &= \int_0^1 \left[\int_0^1 x dx + \int_1^{2-x} (2-x) dx \right] dz \\ &= \int_0^1 \left[\frac{1}{2} - \frac{z^2}{2} + \frac{1}{2} \right] dz \\ &= \int_0^1 1 - \frac{z^2}{2} dz \\ &= 5/6 \end{aligned}$$

Una forma alternativa de calcular esta probabilidad es hacer variar a Z_2 en todo el rango $[0, 2]$ y luego hacer variar a U_3 en todo el rango $[0, 2 - Z_2]$:

$$\begin{aligned}
P(Z_3 \leq 2) &= P(Z_2 \leq 2 - U_3) \\
&= \int_0^2 f_{Z_2}(x) \int_0^{2-x} f_{U_3}(y) dy dx \\
&= \int_0^1 f_{Z_2}(x) \int_0^1 f_{U_3}(y) dy dx + \int_1^2 f_{Z_2}(x) \int_0^{2-x} f_{U_3}(y) dy dx \quad \{\text{Por } \star\} \\
&= \int_0^1 x \int_0^1 1 dy dx + \int_1^2 (2-x) \int_0^{2-x} 1 dy dx \\
&= \int_0^1 x dx + \int_1^2 (2-x)^2 dx \\
&= \frac{1}{2} + \frac{1}{3} \\
&= \frac{5}{6}
\end{aligned}$$

(\star) Si Z_2 varía en $[0, 1]$, $Z \leq 2 - U_3$ para cualquier valor de U_3 , con lo cual hacemos a U_3 variar en todo el rango $[0, 1]$. Pero si Z_2 varía en $[1, 2]$, entonces U_3 solo puede tomar valores en $[0, 2 - Z_2]$.

$$\begin{aligned}
\therefore P(\text{Ganar}) &= P(Z_2 \leq 2) \cdot \frac{1}{3} + P(Z_3 \leq 2) \cdot \frac{2}{3} \\
&= \frac{1}{3} + \frac{5}{6} \cdot \frac{2}{3} \\
&= 0.8
\end{aligned}$$

b) Implementar un algoritmo en computadora que estime la probabilidad de acertar, esto es, la fracción de veces que se acierta en n realizaciones del juego.

En Python, una implementación posible es:

```

from random import uniform

def sim_3(n):
    wins = 0

    # S = summation of uniform r.vs
    def S(k):
        return sum( [uniform(0, 1) for _ in range(k)] )

    for _ in range(n):

```

```
u = uniform(0, 1)
X = S(2) if u < 1/3 else S(3)
wins = wins + 1 if X <= 2 else wins

return wins/n
```

(4) Un supermercado posee 3 cajas. Por una cuestión de ubicación, el 40% de los clientes eligen la caja 1 para pagar, el 32% la caja 2, y el 28% la caja 3.

El tiempo que espera una persona para ser atendida en cada caja distribuye exponencial con medias de 3, 4 y 5 minutos respectivamente.

(a) ¿Cuál es la probabilidad de que un cliente espere menos de 4 minutos para ser atendido?

Recordemos que la media de la exponencial es $1/\lambda$, es decir que $\lambda_1 = 1/3$, $\lambda_2 = 1/4$, $\lambda_3 = 1/5$.

El evento "un cliente espera menos de 4 minutos en ser atendido" se divide en un sub-eventos correspondiente por cada caja. Más formalmente, si T es el tiempo de espera de un cliente arbitrario, y C_i el evento de que el cliente elige la i -ésima caja, entonces por ley de probabilidad total:

$$P(T \leq 4) = \sum_{i=1}^3 P(C_i)P(T \leq 4 \mid C_i) = \sum_{i=1}^3 P(C_i)P(X_i \leq 4)$$

donde X_i es una exponencial con parámetro λ_i . En general,

$$\int_0^4 \lambda_i e^{-\lambda_i x} dx = 1 - e^{-4\lambda_i}$$

de lo cual se sigue:

$$P(X_i \leq 4) = \begin{cases} 0.736 & i = 1 \\ 0.632 & i = 2 \\ 0.55 & i = 3 \end{cases}$$

$$\therefore P(T \leq 4) = 0.40 \cdot 0.736 + 0.32 \cdot 0.632 + 0.28 \cdot 0.55 = 0.65064$$

(b) Si el cliente tuvo que esperar más de 4 minutos. ¿Cuál es la probabilidad de que el cliente haya elegido cada una de las cajas?

Sale fácil con teorema de Bayes y los resultados de (a):

$$P(C_i \mid T \leq 4) = \frac{P(T \leq 4 \mid C_i)P(C_i)}{P(T \leq 4)}$$

(c) Simule el problema y estime las probabilidades anteriores con 1000 iteraciones.

An ugly Python implementation:

```
from numpy.random import exponential
from random import uniform

def sim_4(n):

    wins = 0
    rates = [1/3, 1/4, 1/5]

    def choose_rate(x):
        if x <= 0.4:
            return 0
        return 1 if u <= 0.4 + 0.32 else 2

    for _ in range(n):

        u = uniform(0, 1)
        rate = rates[choose_rate(u)]

        e = exponential(1/rate)

        wins = wins + 1 if e <= 4 else wins

    return wins/n
```

(5) Calcule con el método de Monte Carlo y escriba simulaciones para cada una de las siguientes integrales.

(a) $\int_0^1 (1-x^2)^{3/2} dx$

Sea $f_U(x)$ la densidad de $U \sim \mathcal{U}(0, 1)$. Entonces

$$\theta := \int_0^1 (1-x^2)^{3/2} dx = \int_0^1 (1-x^2)^{3/2} f_U(x) dx = \mathbb{E}[g(U)]$$

con $g(x) = (1-x^2)^{3/2}$. Por lo tanto, $\theta \approx \frac{1}{n} \sum_{i=1}^n g(u_i)$ con u_1, \dots, u_n realizaciones de $U_1, \dots, U_n \sim \mathcal{U}(0, 1)$. Es decir,

$$\theta \approx \frac{1}{n} \sum_{i=1}^n (1-u_i^2)^{3/2}$$

En Python,

```
from random import uniform
def sim_5a(n):
    uniforms = [uniform(0, 1) for _ in range(n)]
    return 1/n * sum( [ (1 - u**2)**1.5 for u in uniforms ] )
```

$$(b) \int_2^3 \frac{x}{x^2-1} dx$$

Si hacemos $u = x - 2$, tal que $du = dx$ y $x = 2 + u$, tenemos

$$\theta := \int_2^3 \frac{x}{x^2-1} dx = \int_0^1 \frac{2+u}{(2+u)^2-1} du = \int_0^1 \frac{2+u}{u^2-4u+3}$$

Entonces $\theta = \int_0^1 g(x)f_U(x) dx = \mathbb{E}[g(U)]$ con $U \sim \mathcal{U}(0,1)$, $g(x) = \frac{2+x}{(2+x)^2-1}$.

$$\therefore \theta \approx \frac{1}{n} \sum_{i=1}^n \frac{2+u_i}{u_i^2-4u_i+3}$$

donde u_1, \dots, u_n son realizaciones de $U_1, \dots, U_n \sim \mathcal{U}(0,1)$ independientes para $n \geq 1$. En Python,

```
def sim_5b(n):
    uniforms = [uniform(0, 1) for _ in range(n)]
    return 1/n * sum( [ (2 + u)/((2+u)**2 - 1) for u in uniforms ] )
```

Con $n = 1000$ obtuve $\theta \approx 0.487$, y el valor real de la integral es $\theta = 0.49$.

$$(c) \int_0^\infty x(1+x^2)^{-2} dx$$

Sea $u := \frac{1}{x+1}$, de manera que

$$du = -\frac{1}{(x+1)^2} dx = -u^2 dx, \quad x = \frac{1-u}{u} = \frac{1}{u} - 1$$

Entonces

$$\theta := \int_0^\infty g(x) dx = - \int_0^1 \frac{g(\frac{1}{u}-1)}{u^2} du = - \int_0^1 \frac{(\frac{1}{u}-1)(1+(\frac{1}{u}-1)^2)^{-2}}{u^2} du$$

Luego

$$\theta = - \int_0^1 \psi(u) f_U(u) du = -\mathbb{E}[\psi(U)]$$

con $\psi(u) = g(\frac{1}{u}-1)/u^2$ y $U \sim \mathcal{U}(0,1)$. En consecuencia,

$$\theta \approx \frac{1}{n} \sum_{i=1}^n \psi(u_i)$$

donde u_1, \dots, u_n son realizaciones de $U_1, \dots, U_n \sim \mathcal{U}(0,1)$ independientes (no confundir con u , la variable de integración) y $n \geq 1$. En Python,

```
def sim_5c(n):
    uniforms = [uniform(0, 1) for _ in range(n)]

    def h(u):
        return (1/u - 1)*(1 + (1/u - 1)**2)**(-2) * (1/u**2)

    return 1/n * sum( [h(u) for u in uniforms] )
```

Con $n = 1000$, la simulación dio $\varphi \approx 0.504$, y el valor real de la integral es $\theta = \frac{1}{2}$.

$$(d) \int_{-\infty}^{\infty} e^{-x^2} dx$$

Let $g(x) := e^{-x^2}$. Clearly, $g(-x) = e^{-(-x)^2} = g(x)$, which means g is symmetric.

$\therefore \int_{-\infty}^{\infty} e^{-x^2} dx = 2 \int_0^{\infty} e^{-x^2} dx$. If we let

$$\psi = \frac{1}{x+1}, \quad d\psi = -\psi^2, \quad x = \frac{1}{\psi} - 1$$

Entonces

$$\theta := 2 \int_0^{\infty} e^{-x^2} dx = 2 \int_0^1 \frac{\exp\left(-\left(\left(\frac{1}{\psi} - 1\right)^2\right)\right)}{\psi^2} f_U(\psi) d\psi$$

donde f_U es la densidad de $U \sim \mathcal{U}(0, 1)$. $\therefore \theta = \mathbb{E}[h(U)]$ con $h(x) = \exp\left(-\left(\left(\frac{1}{x} - 1\right)^2\right)\right) / x^2$.

$$\therefore \theta \approx \frac{1}{n} \sum_{i=1}^n \frac{\exp\left(-\left(\left(\frac{1}{u_i} - 1\right)^2\right)\right)}{u_i^2}, \quad n \geq 1$$

con u_1, \dots, u_n realizaciones de $U_1, \dots, U_n \sim \mathcal{U}(0, 1)$. En Python,

```
from math import exp
from random import uniform

def sim_5d(n):
    def h(u):
        return exp(-( ( 1/u - 1 )**2)) / (u**2)

    uniforms = [uniform(0, 1) for _ in range(n)]
    return 2*sum([ h(u) for u in uniforms ]) / n
```

La simulación con $n = 1000$ dio $\theta \approx 1.73$ que es casi exactamente $\sqrt{\pi}$, el verdadero valor.

Nota. El ejercicio se puede resolver aplicando directamente la sustitución sin notar que g es simétrica, aunque se dan pasos previos algo más engorrosos:

$$\begin{aligned}
\int_{-\infty}^{\infty} g(x) \, dx &= \lim_{t \rightarrow -\infty} \int_t^0 g(x) \, dx + \lim_{t \rightarrow \infty} \int_0^{\infty} g(x) \, dx \\
&= - \int_{\lim_{t \rightarrow -\infty} \psi(t)}^0 \frac{g(\frac{1}{\psi} - 1)}{\psi^2} \, d\psi - \int_0^{\lim_{t \rightarrow \infty} \psi(t)} \frac{g(\frac{1}{\psi} - 1)}{\psi^2} \, d\psi \\
&= \int_0^1 \frac{g(\frac{1}{\psi} - 1)}{\psi^2} \, d\psi + \int_0^1 \frac{g(\frac{1}{\psi} - 1)}{\psi^2} \, d\psi \\
&= 2 \int_0^1 \frac{g(\frac{1}{\psi} - 1)}{\psi^2} \, d\psi
\end{aligned}$$

A partir de acá todo es igual.

$$(e) \theta := \int_0^1 \int_0^1 \exp((x+y)^2) dx dy$$

Let $X, Y \sim \mathcal{U}(0, 1)$ independent with densities f_X, f_Y . Clearly, their joint density function is

$$f_{X,Y}(x, y) = f_X(x)f_Y(y) = \mathcal{I}_{(0,1) \times (0,1)}$$

Then

$$\theta = \int_0^1 \int_0^1 g(x, y) f_{X,Y}(x, y) dx dy = \mathbb{E}[g(X, Y)]$$

$$\therefore \theta \approx \frac{1}{n} \sum_{i=1}^n g(x_i, y_i), \quad n \geq 1$$

where $(x_1, y_1), \dots, (x_n, y_n)$ are realizations of $(X_1, Y_1), \dots, (X_n, Y_n)$, with $X_i, Y_i \sim \mathcal{U}(0, 1)$ for $1 \leq i \leq n$.

In Python,

```
from math import exp
from random import uniform

def sim_5e(n):
    def g(x, y):
        return exp((x+y)**2)

    X = [uniform(0, 1) for _ in range(n)]
    Y = [uniform(0, 1) for _ in range(n)]

    return sum([ g(x, y) for x, y in zip(X, Y) ]) / n
```

With $n = 100000$, the estimation was 4.902, and the true value of the integral according to Wolfram Alpha is 4.899.

(7) Definamos

$$N = \min_n \left\{ n : \sum_{i=1}^n U_i > 1 \right\}, \quad U_i \sim \mathcal{U}(0, 1)$$

- (a) Estimar $\mathbb{E}[N]$ generating n values for N .
(b) Calculate $E[N]$.

Notemos que $Im(N) = \{2, 3, \dots\}$. Calculemos $p_N(n)$ la función de probabilidad de masa de N . Para ello, observemos que

$$P(N = n) = P\left(\sum_{i=1}^{n-1} U_i \leq 1 \cap U_n > 1 - \sum_{i=1}^{n-1} U_i\right)$$

Sea $X(n) = \sum_{i=1}^n U_i$ y, para simplificar la notación, usemos $\varphi(n, x) := P(X(n) \leq x)$ para $x \geq 0$. Es decir, $\varphi(n, x)$ es la función de probabilidad acumulada de $X(n)$. Entonces, la expresión de arriba nos queda

$$P(N = n) = P(X(n-1) < 1 \cap U_n > 1 - X(n-1))$$

De acuerdo con la ley de probabilidad total,

$$\begin{aligned} \varphi(n, x) &= P(X(n) \leq x) \\ &= P(U_n \leq x - X(n-1)) \\ &= \int_{\mathbb{R}} P(U_n \leq x - z) P(X(n-1) = z) \, dz \\ &= \int_0^x (x - z) \frac{d}{dz} \varphi(n-1, z) \, dz \end{aligned}$$

donde estamos utilizando el hecho de que la función de densidad es la derivada de la función de probabilidad acumulada. Esta expresión recursiva de $\varphi(n, x)$ nos permite avanzar: si restringimos $x \in [0, 1]$,

$$\begin{aligned} \varphi(2, x) &= \int_0^x (x - z) \frac{d}{dz} \varphi(1, z) \, dz \\ &= \int_0^x (x - z) \, dz \\ &= \frac{x^2}{2} \end{aligned}$$

Siguiendo probando casos, obtenemos

$$\begin{aligned}
 f(x, 3) &= \int_0^x (x - z) z \, dz = \frac{x^3}{6} \\
 f(x, 4) &= \int_0^x (x - z) \frac{z^2}{2} \, dz = \frac{x^4}{24} \\
 &\vdots \\
 f(x, n) &= (x - t) \frac{d}{dz} \varphi(x, n - 1)(x := z) \, dz = \frac{x^n}{n!}
 \end{aligned}$$

Es decir, obtenemos una expresión general para la probabilidad acumulada de $X(n)$, de lo cual se deduce que la densidad de $X(n)$ es

$$\psi(n, x) = \frac{d}{dx} \frac{x^n}{n!} = \frac{nx^{n-1}}{n!} = \frac{x^{n-1}}{(n-1)!} = \varphi(n-1, x), \quad 0 \leq x \leq 1$$

Por lo tanto,

$$\begin{aligned}
 P(N = n) &= P(X(n-1) < 1 \cap U_n > 1 - X(n-1)) \\
 &= \int_0^1 \psi(n-1, x) P(U_n > 1 - x) \, dx \\
 &= \int_0^1 \frac{x^{n-2}}{(n-2)!} (1 - P(U_n \leq 1 - x)) \, dx \\
 &= \frac{1}{(n-2)!} \int_0^1 x^{n-1} \, dx \\
 &= \frac{1}{(n-2)!} \frac{1^n}{n} \\
 &= \frac{1}{n(n-2)!}
 \end{aligned}$$

Por lo tanto,

$$\mathbb{E}[N] = \sum_{i=2}^{\infty} \frac{1}{(i-2)!} = \sum_{i=0}^{\infty} \frac{1}{i!} = e$$

La simulación está de acuerdo, porque en Python

```
def sim_7(iterations):  
  
    counts = []  
    n, s = 0, 0  
  
    for _ in range(iterations):  
        while s <= 1:  
            n += 1  
            s += uniform(0, 1)  
        counts.append(n)  
        n, s = 0, 0  
  
    return sum(counts)/len(counts)
```

nos devuelve siempre aproximadamente e .

(8) Sea

$$N = \max_n \left(\prod_{i=1}^n U_i \geq e^{-3} \right)$$

con $\prod_{i=1}^0 U_i = 1$. Estimar mediante simulaciones $E[N]$ y $P(N = n)$ para $0 \leq n \leq 6$.

Solución. Estimar con simulaciones es aburrido y fácil. Más vale pensemos en el problema y tratemos de dar p_N , la función de probabilidad de masa de N .

Claramente,

$$\begin{aligned} P(N = n) &= P \left(\prod_{i=1}^{n-1} U_i \geq e^{-3} \cap U_n \cdot \prod_{i=1}^{n-1} U_i < e^{-3} \right) \\ &= P \left(\prod_{i=1}^{n-1} U_i \geq e^{-3} \cap U_n < \frac{e^{-3}}{\prod_{i=1}^{n-1} U_i} \right) \end{aligned}$$

Para simplificar la notación, hagamos $X(n) := \prod_{i=1}^n U_i$ y $\varphi(n, x) = P(X(n) \leq x)$ para $x \in [0, 1]$. Es decir, φ es la CDF de $X(n)$. Asumimos que $X(0) = 1$, con lo cual $\varphi(0, x) = 1$. Es fácil ver que $\varphi(1, x) = x$. Veamos además que

$$\begin{aligned} \varphi(2, x) &= P(U_1 U_2 \leq x) \\ &= P(U_1 \leq \frac{x}{U_2}) \\ &= \int_0^1 P(U_1 \leq \frac{x}{z}) P(U_2 = z) dz \\ &= \int_0^1 P \left(U_1 \leq \frac{x}{z} \right) dz \end{aligned}$$

Como $x/z \leq 1$ para $z \in [x, 1]$ y $x/z > 1$ para $z \in [0, x)$,

$$\begin{aligned}
\int_0^1 P\left(U_1 \leq \frac{x}{z}\right) dz &= \int_0^x P\left(U_1 \leq \frac{x}{z}\right) dz + \int_x^1 P\left(U_1 \leq \frac{x}{z}\right) dz \\
&= \int_0^x 1 dz + \int_x^1 \frac{x}{z} dz \\
&= x + x \left[\ln z \right]_x^1 \\
&= x - x \ln x
\end{aligned}$$

Una forma alternativa de dar con la probabilidad buscada es la siguiente:

$$\begin{aligned}
\varphi(2, x) &= P(U_1 U_2 \leq x) \\
&= \int_0^x \int_0^1 dy dz + \int_x^1 \int_0^{x/z} dy dz \\
&= \int_0^x dz + \int_x^1 \frac{x}{z} dz \\
&= x + x \int_x^1 \frac{1}{z} dz \\
&= x + x \left[\ln z \right]_x^1 \\
&= x + x(\ln 1 - \ln x) \\
&= x - x \ln x
\end{aligned}$$

La forma dada antes, usando la ley de probabilidad total, es más complicada para este caso particular, pero más útil para el caso general.

Generalizando, llegamos a

$$\begin{aligned}
\varphi(n, x) &= P\left(\prod_{i=1}^{n-1} U_i \leq \frac{x}{U_n}\right) \\
&= \int_0^1 \varphi\left(n-1, \frac{x}{z}\right) P(U_n = z) dz \\
&= \int_0^1 \varphi\left(n-1, \frac{x}{z}\right) dz \\
&= \int_0^x \varphi\left(n-1, \frac{x}{z}\right) dz + \int_x^1 \varphi\left(n-1, \frac{x}{z}\right) dz \\
&= \int_0^x 1 dz + \int_x^1 \varphi\left(n-1, \frac{x}{z}\right) dz \\
&= x + \int_x^1 \varphi\left(n-1, \frac{x}{z}\right) dz
\end{aligned}$$

Entonces, usando el hecho de que conocemos $\varphi(2, x)$, tenemos

$$\begin{aligned}
\varphi(3, x) &= x + \int_x^1 \varphi\left(2, \frac{x}{z}\right) dz \\
&= x + \int_x^1 \frac{x}{z} - \frac{x}{z} \ln \frac{x}{z} dz \\
&= x - x \ln x + \frac{x}{2} \ln^2 x
\end{aligned}$$

$$\begin{aligned}
\varphi(4, x) &= x + \int_x^1 \varphi\left(3, \frac{x}{z}\right) dz \\
&= x + \int_x^1 \frac{x}{z} - \frac{x}{z} \ln\left(\frac{x}{z}\right) + \frac{x}{2z} \ln^2\left(\frac{x}{z}\right) dz \\
&= x - x \ln x + \frac{x \ln^2 x}{2} - \frac{x \ln^3 x}{6}
\end{aligned}$$

Visto estos resultados, notamos (aunque no demostramos formalmente) que:

$$\varphi(n, x) = \sum_{k=0}^{n-1} (-1)^k \frac{x \ln^k x}{k!}$$

Por lo tanto, la función de probabilidad de masa del producto de n uniformes aleatorias es

$$\begin{aligned}
\psi(n, x) &= \frac{d}{dx} \varphi(n, x) \\
&= \sum_{k=0}^{n-1} (-1)^k \frac{\ln^k x + k \ln^{k-1} x}{k!} \\
&= 1 - (\ln x + 1) + \left(\ln x + \frac{\ln^2 x}{2} \right) - \left(\frac{\ln^3 x}{6} + \frac{1}{2} \ln^2 x \right) + \dots \\
&= (-1)^{n-1} \frac{\ln^{n-1} x}{(n-1)!} \quad \text{\{Suma telescópica\}}
\end{aligned}$$

Luego

$$\begin{aligned}
P(N = n) &= P\left(\prod_{i=1}^{n-1} U_i \geq e^{-3} \cap U_n \cdot \prod_{i=1}^{n-1} U_i < e^{-3}\right) \\
&= P\left(\prod_{i=1}^{n-1} U_i \geq e^{-3} \cap U_n < \frac{e^{-3}}{\prod_{i=1}^{n-1} U_i < e^{-3}}\right) \\
&= \int_{e^{-3}}^1 P(U_n < \frac{e^{-3}}{z}) P\left(\prod_{i=1}^{n-1} U_i = z\right) dz \\
&= \int_{e^{-3}}^1 \left(\frac{e^{-3}}{z}\right) \psi(n-1, z) dz \\
&= \int_{e^{-3}}^1 \left(\frac{e^{-3}}{z}\right) (-1)^{n-2} \frac{\ln^{n-2} z}{(n-2)!} dz \\
&= \frac{(-1)^{n-2} e^{-3}}{(n-2)!} \int_{e^{-3}}^1 \frac{\ln^{n-2} z}{z} dz \\
&= \frac{(-1)^{n-2} e^{-3}}{(n-2)!} \left(-\frac{(-3)^{n-1}}{n-1}\right) \\
&= \frac{3^{n-1} e^{-3}}{(n-1)!} \\
&= \frac{3^{n-1}}{e^3 (n-1)!}
\end{aligned}$$

Por lo tanto,

$$\begin{aligned}
\mathbb{E}[N] &= \sum_{k=1}^{\infty} \frac{k 3^{k-1}}{e^3 (k-1)!} \\
&= \frac{1}{e^3} \sum_{k=0}^{\infty} 3^k \frac{k+1}{k!} \\
&= \frac{1}{e^3} \left(\sum_{k=0}^{\infty} k \frac{3^k}{k!} + \sum_{k=0}^{\infty} \frac{3^k}{k!} \right) \\
&= \frac{1}{e^3} \left(\frac{1}{e^{-3}} \sum_{k=0}^{\infty} k \cdot e^{-3} \frac{3^k}{k!} + \sum_{k=0}^{\infty} \frac{3^k}{k!} \right)
\end{aligned}$$

Ahora somos pillos y vemos que la primera sumatoria es el valor esperado de una Poisson con parámetro $\lambda = 3$, y por lo tanto es 3. la otra sumatoria es e^3 , dado que $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$. Como $1/e^{-3} = e^3$,

$$\therefore \mathbb{E}[N] = \frac{1}{e^3} (3e^3 + e^3) = \frac{1}{e^3} \cdot 4e^3 = 4$$

Este valor coincide con la estimación de la siguiente simulación:

```
def sim_9(iterations):

    counts = []
    n, s = 0, 1

    for _ in range(iterations):
        while s >= exp(-3):
            n += 1
            s *= uniform(0, 1)
        counts.append(n)
        n, s = 0, 1

    return sum(counts)/len(counts)
```

(9) Un juego consiste en dos pasos. En el primer paso se tira un dado convencional. Si sale 1 o 6 tira un nuevo dado y se le otorga al jugador como puntaje el doble del resultado obtenido en esta nueva tirada; pero si sale 2, 3, 4 o 5 en la primer tirada, el jugador debería tirar dos nuevos dados, y recibiría como puntaje la suma de los dados. Si el puntaje del jugador excede los 6 puntos entonces gana.

(a) Calcular la probabilidad de que un jugador gane.

Sea $X = 1$ si el primer dado es 1 o 6, $X = 0$ si el primer dado es 2, 3, 4 o 5. Claramente X es una Bernoulli con $p = 1/3$. Entonces, de acuerdo con la ley de probabilidad total,

$$P(\text{Ganar}) = P(2U > 6 \mid X = 1)P(X = 1) + P(U_1 + U_2 > 6 \mid X = 0)P(X = 0)$$

con U, U_2, U_3 uniformes discretas en $\{1, \dots, 6\}$. Claramente, $U_1 + U_2 > 6$ y $2U$ son independientes de X . Por lo tanto

$$P(\text{Ganar}) = P(U > 3)1/3 + P(U_1 + U_2 > 6)2/3$$

Ahora bien, $P(U_1 + U_2 > 6) = 1 - P(U_1 \leq 6 - U_2)$, y

$$\begin{aligned} P(U_1 \leq 6 - U_2) &= \sum_{k=1}^6 P(U_1 \leq 6 - k)P(U_2 = k) \\ &= \sum_{k=1}^6 \frac{6 - k}{6} \frac{1}{6} \\ &= \frac{1}{6} \sum_{k=1}^6 1 - \frac{k}{6} \\ &= 0.416 \end{aligned}$$

Por lo tanto, $P(U_1 + U_2 > 6) = 0.583$. Es fácil ver que $P(U > 3) = \frac{1}{2}$. Luego

$$P(\text{Ganar}) = \frac{1}{2} \cdot \frac{1}{3} + 0.583 \cdot \frac{2}{3} = 0.555$$

(b) Estimar mediante simulación.


```
from random import randint
def sim_8(iterations):

    wins = 0

    for _ in range(iterations):
        first_dice = randint(1, 6)
        X = first_dice == 0 or first_dice == 6

        if X:
            result = 2*randint(1, 6)
        else:
            result = randint(1, 6) + randint(1, 6)

        wins = wins + 1 if result > 6 else wins

    return wins/iterations
```

(Extra: 1) Estimar $\int_{-1}^{\infty} x^3 e^{-x^3} dx$.

Sea $u(x) = \frac{1}{2+x}$, tal que $\lim_{x \rightarrow \infty} u(x) = 0$ y $u(-1) = 1$.
Observemos que

$$du = -\frac{1}{(2+x)^2} dx = -u^2 dx, \quad x = \frac{1}{u} - 2$$

Entonces

$$\begin{aligned} \int_{-1}^{\infty} x^3 \exp(-x^3) dx &= \int_0^1 \frac{\left(\frac{1}{u} - 2\right)^3 \exp\left(-\left(\frac{1}{u} - 2\right)^3\right)}{u^2} du \\ &= \int_0^1 \varphi(u) f_U(u) du \end{aligned}$$

con $\varphi(u)$ la horripilancia que queda en la integral y f_U la densidad de una uniforme en $(0, 1)$. Luego estimamos la integral como

$$\frac{1}{N} \sum_{i=1}^N \varphi(u_i)$$

Deseamos una sustitución $u(x)$ tal que $u(-1) = 0$ y $\lim_{x \rightarrow \infty} u(x) = 1$. Es claro que

$$u := 1 + \frac{1}{x}$$

satisface esta propiedad, y que

$$du = -\frac{1}{x^2} dx = -(u-1)^2 dx, \quad x = \frac{1}{u-1}$$

Por lo tanto,

$$\begin{aligned} \int_{-1}^{\infty} x^3 \exp(-x^3) dx &= \int_0^1 \frac{\exp\left(-\frac{1}{(u-1)^3}\right)}{-(u-1)^2(u-1)^3} du \\ &= - \int_0^1 \frac{\exp\left(-\frac{1}{(u-1)^3}\right)}{(u-1)^5} du \\ &\approx \frac{1}{N} \sum_{i=1}^N \varphi(u_i) \end{aligned}$$

con $N \geq 1$ y

$$\varphi(u) = \frac{\exp\left(-\frac{1}{(u-1)^3}\right)}{(u-1)^3} du$$