Pontificia Universidad Católica de Chile
Escuela de Ingeniería
Departamento de Ingeniería Industrial y de Sistemas
ICS2123 – Modelos Estocásticos
Profesores Maximiliano González, Nicolás Vargas y Andrés Navarro
Ayudante Jefe Francisco Tamarín
Segundo Semestre 2024

Tarea 1:
Repaso de Probabilidades

Integrantes

- **Integrante 1:** Jacinta Ortiz (Sección 4)
- **Integrante 2:** Vicente Lavagnino (Sección 2)

Instrucciones

- Esta tarea debe ser realizada **individualmente** o **en parejas** (pueden ser de diferentes secciones dentro de las que son coordinadas). Deben indicar en la primera plana del PDF los nombres de los/as integrantes, junto con las secciones respectivas de c/u.
- Tienen plazo hasta el **Martes 27 de agosto a las 23:59 hrs** para entregar sus soluciones.
- Deben entregar un archivo reporte, **en formato PDF**. En el archivo PDF deben entregar sus respuestas de los 3 problemas. Los códigos implementados deben adjuntarse como un archivo .zip aparte en el buzón. En el PDF deberá estar el desarrollo, resumen de resultados y correspondiente análisis. Se recomienda fuertemente utilizar LaTeX. Sin embargo, estas también pueden ser escritas en algún otro editor de texto o a mano y escaneadas, pero **debe estar ordenado y legible**, de lo contrario, se descontarán 3 décimas a la nota final de la tarea. Cualquier código que utilicen de internet debe ser debidamente citado insertando el *link* asociado.
- Si realizan la tarea completa en LaTeX, incluyendo todas las fórmulas y ecuaciones, tendrán **3 décimas extras**.
- La tarea tiene un total de 60 puntos. La nota final será calculada como la suma de las tres preguntas mediante la fórmula
$$Nota = \frac{P_1 + P_2 + P_3}{60} \cdot 6 + 1$$
Donde P_i corresponde al puntaje de la pregunta i. Se aproxima a dos decimales.
- Las preguntas se harán a través del foro de discusión asociado a la Tarea 1 en Canvas, por lo que les pedimos que **NO** manden dudas por mail sobre la tarea. De esta manera, todos tienen acceso a las respuestas.
- Seremos estrictos en que **no se aceptarán envíos por mail**. Evite problemas de internet, del sistema Canvas, de la calidad de la foto o cualquier otro que pudiese ocurrir al entregar sobre la hora. El/La alumno/a que envíe su tarea después de la hora límite, tendrá **nota 1.0**.

Problema 1 (20 puntos)

Parte 1 (10 puntos)

a)

Si A, B, y C son eventos mutuamente independientes, entonces A y $B \cup C$ son independientes, y $A \cap \bar{B}$ y C también son independientes.

Esto se demuestra de la siguiente manera:

Notemos que para dos eventos (A,B) mutuamente independientes se cumple que :

$$P(A \cap B) = P(A) \cdot P(B) \tag{1}$$

$$P(A|B) = P(A) \tag{2}$$

Se tiene que,

$$P(A \cap (B \cup C)) =$$
$$P(A \cap B) \cup P(A \cap C) =$$
$$P(A \cap B) + P(A \cap C) - P((A \cup B) \cup (A \cap C)) =$$
$$P(A)P(B) + P(A)P(C) - P(A)^2 P(B)P(C) =$$
$$P(A) \cdot (P(B) + P(C) - P(B)P(C))$$

Lo cual cumple con la condición (1) y queda demostrado que A y $B \cup C$ son independientes.

Por otro lado,

$$P(A \cap \bar{B}) = P(A) \cdot (1 - P(B)) = P(A) - P(A) \cdot P(B) \tag{3}$$

$$P(A \cap \bar{B}|C) = \frac{(P(A) - P(A)P(B)) \cap P(C)}{P(C)} = \frac{(P(A)P(C) - P(A)P(B)P(C)}{P(C)} = P(A) - P(A)P(B) \tag{4}$$

Dado que esto cumple la lógica de la ecuacion (2) queda demostrado que $A \cap \bar{B}$ y C son independientes.

b)

Si $P(A|C) \geq P(B|C)$ y $P(A|\bar{C}) \geq P(B|\bar{C})$. Entonces se tiene que con la primera desigualdad:

$$\frac{P(A \cap C)}{P(C)} \geq \frac{P(B \cap C)}{P(C)} \tag{5}$$

Entonces,

$$P(A \cap C) \geq P(B \cap C) \tag{6}$$

Siguiendo la misma lógica con la segunda desigualdad,

$$P(A \cap \bar{C}) \geq P(B \cap \bar{C}) \tag{7}$$

Es decir,

$$P(A) - P(A)P(C) \geq P(B) - P(B)P(C) \tag{8}$$

$$P(B \cap C) \geq P(B) - P(A) + P(A \cap C) \tag{9}$$

Ahora con la desigualdad (7) podemos reemplazar $P(B \cap C)$ por $P(A \cap C)$ en la ecuación (9):

$$P(A \cap C) \geq P(B) - P(A) + P(A \cap C) \tag{10}$$

Lo que resulta:

$$P(A) \geq P(B) \tag{11}$$

Quedando demostrado que P(A) $\geq$ P(B).

c)

Queremos demostrar que:

$$P(A) \geq \sum_{j=1}^{n} P(A_j) - (n-1) \tag{12}$$

Por un lado tenemos que $P(A)$ y $\sum_{j=1}^{n} P(A_j) \in (0,1)$, así que desde n=2 en adelante $(n-1) < 0$ por lo que se demuestra la hipótesis.
Ahora debemos demostrarlo para n=1. Como tenemos que:

$$\bigcap_{j=1}^{n} A_J \subseteq A \tag{13}$$

Podemos decir que:

$$P(A) \geq \bigcap_{j=1}^{n} A_j \tag{14}$$

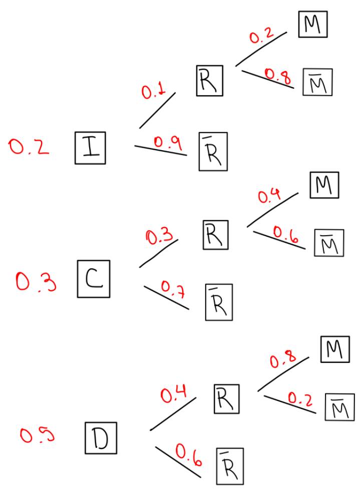
Es decir que, para n=1, $P(A) \geq P(A_1)$. Además, la desigualdad (13) en n=1:

$$P(A) \geq P(A_1) - (1-1) = P(A_1) \tag{15}$$

Quedando demostrado para n=1.

Parte 2 (10 puntos)

Para el problema se realizó el siguiente desarrollo del árbol de probabilidades. Donde I, C, D son los partidos políticos, R significa que se obtiene respuesta, $\bar{R}$ que no se obtiene respuesta, M que se obtiene aporte monetario y $\bar{M}$ que no se obtiene aporte monetario.



d)

Dado que se selecciona <u>un</u> mail al azar, es que la probabilidad de que sólo D sea sometido a juicio, es la misma que decir "la probabilidad de que se seleccione un mail de D y que este haya cometido delito".
Por lo tanto, se requiere que se obtenga una respuesta y además se reciba aporte monetario:

$P[$ sólo D sea sometido a juicio $] = [M|R \cap D] = 0.5 \cdot 0.4 \cdot 0.8 = 0.16.$

e)

Para encontrar la probabilidad de que ningún partido sea sometido a juicio se requiere:

$P[$ (D no sea sometido a juicio) $\cup$ (C no sea sometido a juicio) $\cup$ (I no sea sometido a juicio) $] =$ $0.5 \cdot (1 - 0.4 \cdot 0.8) + 0.3(1 - 0.3 \cdot 0.4) + 0.2(1 - 0.1 \cdot 0.2) = 0.8$

f)

Se pide:

$$P\Big[(D \text{ cometió delito} \mid D \text{ obtuvo respuesta}) \cup (C \text{ cometió delito} \mid C \text{ obtuvo respuesta}) \cup (I \text{ cometió delito} \mid I \text{ obtuvo respuesta})\Big]$$

$$=$$

$$\frac{0.2 \cdot 0.4 \cdot 0.5 + 0.6 \cdot 0.3 \cdot 0.3 + 0.8 \cdot 0.1 \cdot 0.2}{0.5 \cdot 0.4 + 0.3 \cdot 0.3 + 0.2 \cdot 0.1} = 0.35$$

Problema 2 (20 puntos)

a)

i)

Tenemos que t: tiempo entre accidientes $\sim Exponencial(\lambda)$ y T: tiempo de observación $\sim Exponencial(\nu)$. Entonces, el número de accidentes Y dependiente del tiempo T , es decir $Y|T \sim Poisson(\lambda T)$.
Ahora para encontrar la distribución de Y independiente de T (tiempo de observacion) , necesitamos la distribución marginal de Y:

$$p_Y(y) = \int_{y \in \Theta(Y,T)} p_{Y|T=t} \cdot f_T(t) dt \tag{16}$$

Es decir, buscamos la siguiente integral:

$$\int_0^\infty \frac{(\lambda t)^y e^{-\lambda t}}{y!} \cdot \nu e^{-\nu t} \, dt$$

Podemos simplificarla de la siguiente manera:

$$\frac{\lambda^y}{y!} \nu \int_0^\infty t^y e^{-(\lambda+\nu)t} \, dt$$

Reconocemos que la integral es una función gamma, que tiene la forma general:

$$\int_0^\infty t^n e^{-bt} \, dt = \frac{\Gamma(n+1)}{b^{n+1}}$$

En este caso, $n = y$ y $b = \lambda + \nu$. Entonces, la integral se puede resolver como:

$$\frac{\lambda^y \nu}{y!} \int_0^\infty t^y e^{-(\lambda+\nu)t} \, dt = \frac{\lambda \nu}{y!} \cdot \frac{\Gamma(y+1)}{(\lambda+\nu)^{y+1}}$$

Sabemos que $\Gamma(y+1) = y!$ cuando y es un número entero. Por lo tanto, la expresión se simplifica a:

$$\frac{\lambda^y \nu}{(\lambda+\nu)^{y+1}}$$

Finalmente, el resultado de la integral es:

$$\int_0^\infty \frac{(\lambda t)^y e^{-\lambda t}}{y!} \cdot \nu e^{-\nu t} \, dt = \frac{\lambda^y \nu}{(\lambda+\nu)^{y+1}}$$

Lo cual se puede reescribir a la distribución de Y a demostrar:

$$\frac{\nu}{(\lambda+\nu)} \cdot (\frac{\lambda}{\lambda+\nu})^y$$

ii)

Para encontrar el valor esperado de Y, utilizamos:

$$E[Y] = \sum_{y=0}^{\infty} y \cdot P(Y = y)$$

Dado que Y sigue una distribución Poisson con parámetro λT, y T sigue una distribución Exponencial con parámetro ν, podemos expresar el valor esperado de Y como:

$$E[Y] = E[E[Y|T]]$$

Como $Y|T = t$ sigue una distribución Poisson(λt), tenemos que:

$$E[Y|T = t] = \lambda t$$
$$E[Y] = E[\lambda T] = \lambda E[T]$$

Dado que T sigue una distribución Exponencial(ν), entonces, el valor esperado de Y es:

$$E[Y] = \frac{\lambda}{\nu}$$

Ahora, calculamos la varianza de Y:

$$Var(Y) = E[Var(Y|T)] + Var(E[Y|T])$$

Primero, calculemos $E[Var(Y|T)]$:

$$E[Var(Y|T)] = E[\lambda T] = \lambda E[T] = \frac{\lambda}{\nu}$$

Ahora, calculemos $Var(E[Y|T])$:

$$Var(E[Y|T]) = Var(\lambda T) = \lambda^2 Var(T)$$

Dado que T sigue una distribución Exponencial(ν):

$$Var(T) = \frac{1}{\nu^2}$$
$$Var(E[Y|T]) = \lambda^2 \cdot \frac{1}{\nu^2} = \frac{\lambda^2}{\nu^2}$$

Sumando se tiene que:

$$Var(Y) = \frac{\lambda}{\nu} + \frac{\lambda^2}{\nu^2}$$

iii)

Queremos demostar que:

$$P(Z = z) = (1 - p)^{z-1} p$$

Por enunciado, sabemos que $Z = z$ es equivalente a $Y = z - 1$. Entonces:

$$P(Z = z) = P(Y = z - 1)$$

Sustituyendo con $P(Y=y)$ demostrado en i):

$$P(Z=z) = \frac{\nu}{\lambda+\nu}\left(\frac{\lambda}{\lambda+\nu}\right)^{z-1}$$

Lo podemos reescribir como:

$$P(Z=z) = \frac{\nu}{\lambda+\nu}(1-\frac{\nu}{\lambda+\nu})^{z-1}$$

Lo cual es claramente una distribución Geométrica de de parámetro p= $\frac{\nu}{\lambda+\nu}$

b)

iv)

Tenemos que X: cantidad de errores cometidos en 1 minuto distribuye Poisson($\frac{1}{\upsilon}$) y T: tiempo de desarrollo, distribuye Gamma (k,υ). Ahora, X_t: cantidad de errores en un tiempo T, distribuye Poisson ($\frac{T}{\upsilon}$).
Se pide $P[T|X_{t=k}]$, entonces buscamos:

$$\frac{f_{T,X_{t=k}}(T,X_t)}{f_T(T)} \tag{17}$$

Es decir,

$$\frac{\frac{\upsilon^k \cdot T^{k-1}e^{-\upsilon T}}{\Gamma(k)}}{\frac{(\frac{T}{\upsilon})^k \cdot e^{-\frac{T}{\upsilon}}}{k!}} \tag{18}$$

Simplificando:

$$\frac{\upsilon^{2k}e^{(-\upsilon T+\frac{T}{\upsilon})}\cdot k}{T} \tag{19}$$

v)

Por el teorema de esperanza iteradas tenemos que:

$$E[X] = E[E(X|Y)] \tag{20}$$

Y nos piden $Cov(T,X_t) = E[T\cdot X_t] - E[T][X_t]$. Por un lado tenemos:

- $E[T] = \frac{k}{\upsilon}$
- $E[X_t|T] = \frac{T}{\upsilon}$

Por (17) podemos reescribir Cov(T, X_t) como:

$$E[T\cdot E[X_t|T]] - E[T]\cdot E[E(X_t|T)] = E[T\cdot\frac{T}{\upsilon}] - E[T]\cdot E[\frac{T}{\upsilon}]$$

$$= \frac{1}{\upsilon}\cdot E[T^2] - \frac{1}{\upsilon}\cdot E[T]^2 = \frac{1}{\upsilon}(E[T^2]-E[T]^2) = \frac{1}{\upsilon}\cdot VAR(T) = \frac{1}{\upsilon}\cdot\frac{k}{\upsilon^2} = \frac{k}{\upsilon^3}$$

c) Pérdida de memoria ***revisar vicho

X es una variable aleatoria continua con soporte $[0,\infty)$ con distribución de probabilidad $f(x)$. Asumiendo que distribuye Exponencial con parámetro υ entonces su distribución de probabilidad acumulada $F(X)$ o $P(X\leq x)$ es:

$$1-e^{-\upsilon x} \tag{21}$$

Y podemos definir a $P(X\geq X)$ como 1- F(x), es decir:

$$e^{-\upsilon x} \tag{22}$$

Ahora, si tiene la propiedad de pérdida de memoria debiera cumplirse que :

$$P(X \geq s+t | X \geq s) = P(X \geq t) \tag{23}$$

La parte izquierda de la desigualdad puede escribirse como:

$$= \frac{P(X \geq t+s, X \geq s)}{P(X \geq s)}$$

Como $t, s > 0$, podemos reescribir lo anterior como:

$$= \frac{P(X \geq t+s)}{P(X \geq s)}$$

Reemplazando con (14), tenemos que:

$$\begin{aligned} &= \frac{e^{-\upsilon(t+s)}}{e^{-\upsilon s}} \\ &= e^{-\upsilon t} \\ &= P(X \geq t) \end{aligned}$$

d) Viejito pascuero

Tenemos que $T_1, T_2, ..., T_n$: tiempo de llegada de cada trineo de la flota i $\sim$ Exponencial (λ_i). Si definimos a Z_i como el tiempo mínimo de llegada del primer trineo de la siguiente flota j, entonces buscamos para todo i que:

$$P(T_1^i < Z_{i+1}) \tag{24}$$

Es decir,

$$\prod_{i=1}^{n-1} P(T_1^i < Z_{i+1}) \tag{25}$$

Finalmente, considerando que ambas variables son Exponenciales de parámetros λ_i y $(\lambda_{i+1}, ..., \lambda_n)$ Tenemos:

$$\prod_{i=1}^{n-1} (\frac{\lambda_i}{\sum_{j=i}^{n-1} \lambda_j})$$

Problema 3 (20 puntos)

Parte 1

a)

Demostración Teórica

Suponiendo que A es la puerta ganadora, B y C son las puertas perdedoras y $X \in A, B, C$ es la puerta elegida inicialmente por el jugador.

La probabilidad de que el jugador elija A inicialmente es $\mathbb{P}(X = A) = \frac{1}{3}$

Mientras que la probabilidad de que el jugador elija una de las puertas perdedoras (B o C) es $\mathbb{P}(X = B \text{ o } X = C) = \frac{2}{3}$

En ese sentido, si el jugador no cambia de puerta, la probabilidad de ganar es simplemente la probabilidad de que haya elegido la puerta inicialmente $\mathbb{P}(\text{Ganar sin cambiar}) = \mathbb{P}(X = A) = \frac{1}{3}$

En cambio, si el jugador decide cambiar de puerta después de que una esté abierta, enfrentamos 3 escenarios

- Caso 1: El jugador elige A inicialmente (probabilidad $\frac{1}{3}$). Si el jugador cambia, perderá, ya que se abrirá una de las puertas B o C, ambas perdedoras, y el jugador cambiará a la puerta con la otra opción incorrecta. Así que la probabilidad de ganar al cambiar en este caso es 0.

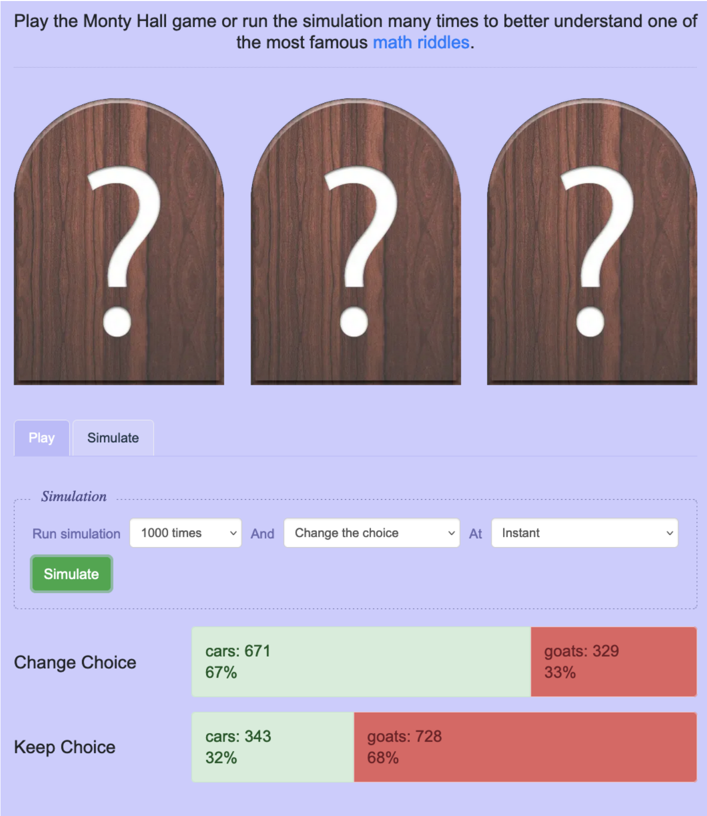
- ◇ Caso 2 :El jugador elige B inicialmente (probabilidad $\frac{1}{3}$). Se abrirá la puerta C, y el jugador cambiará a la puerta A, que tiene la opción ganadora. La probabilidad de ganar al cambiar en este caso es 1.
- ◇ Caso 3: El jugador elige C inicialmente (probabilidad $\frac{1}{3}$). Se abrirá la puerta B, y el jugador cambiará a la puerta A, que tiene el coche. La probabilidad de ganar al cambiar en este caso es 1.

La probabilidad total de ganar al cambiar de puerta es la suma de las probabilidades ponderadas de estos tres casos $\mathbb{P}(\text{Ganar cambiando}) = \left(\frac{1}{3} \times 0\right) + \left(\frac{1}{3} \times 1\right) + \left(\frac{1}{3} \times 1\right) = \frac{2}{3}$

Por lo tanto, la probabilidad de ganar al cambiar de puerta es $\mathbb{P}(\text{Ganar cambiando}) = \frac{2}{3}$ comparada con la probabilidad de ganar sin cambiar, que es solo $\frac{1}{3}$.

Demostración Empírica

Para testear este caso en reiteradas ocasiones, usaremos la siguiente página web `https://www.mathwarehouse.com/monty-hall-simulation-online/`



De esta forma vemos que si hacemos la prueba 1000 veces, aproximadamente 2/3 de las veces acertaremos si cambiamos de de decisión.

Por otra parte, si hacemos nuevamente el ejercicio 1000 veces, sin cambiar nuestra elección, nos podremos dar cuenta de que aproximadamente 1/3 de las veces acertamos.

b)

Demostración Teórica

Suponiendo que hay n personas en un grupo y 365 días en un año. Tenemos que encontrar la probabilidad de que al menos dos personas compartan el mismo cumpleaños.

Sabemos que la probabilidad de que todos tengan cumpleaños diferentes es:

$$P(\text{todos diferentes}) = \frac{365}{365} \cdot \frac{364}{365} \cdot \dots \cdot \frac{365-n+1}{365} = \frac{365!}{(365-n)! \cdot 365^n}$$

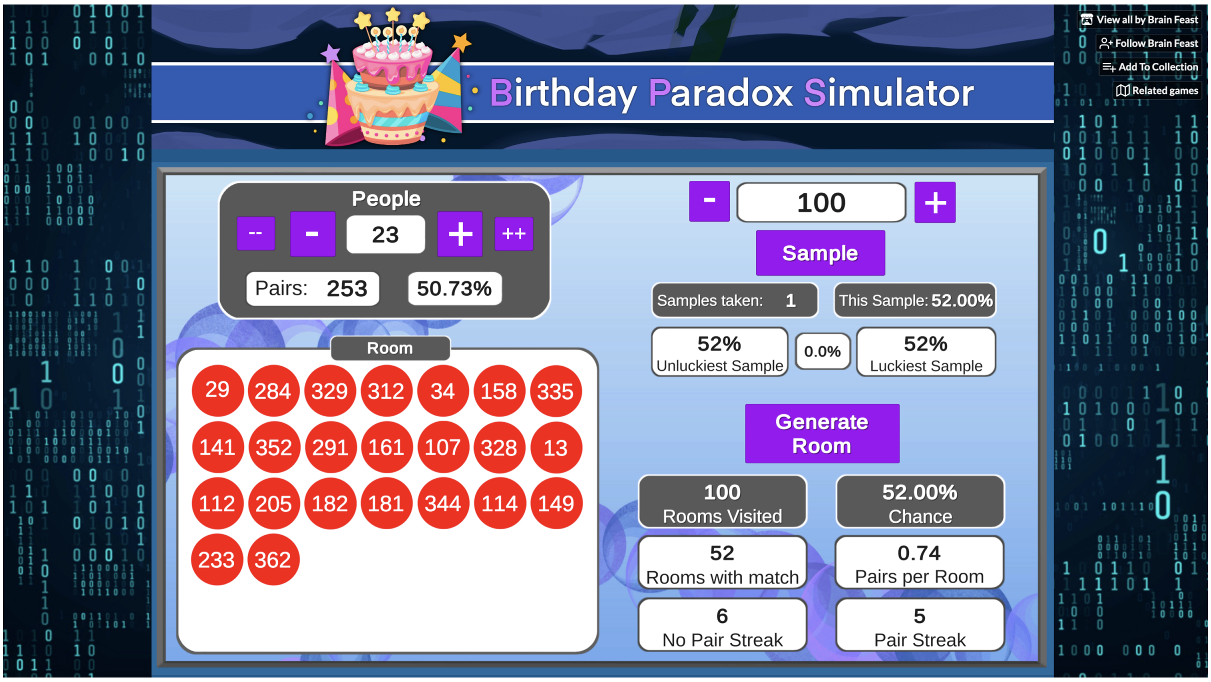
Por lo tanto, la probabilidad de al menos una coincidencia es:

$$P(\text{al menos una coincidencia}) = 1 - \frac{365!}{(365-n)! \cdot 365^n}$$

Demostración Empírica

Para probar de manera práctica los datos usaremos el sistema recreado en `https://brain-feast.itch.io/birthday-paradox-simulator`.

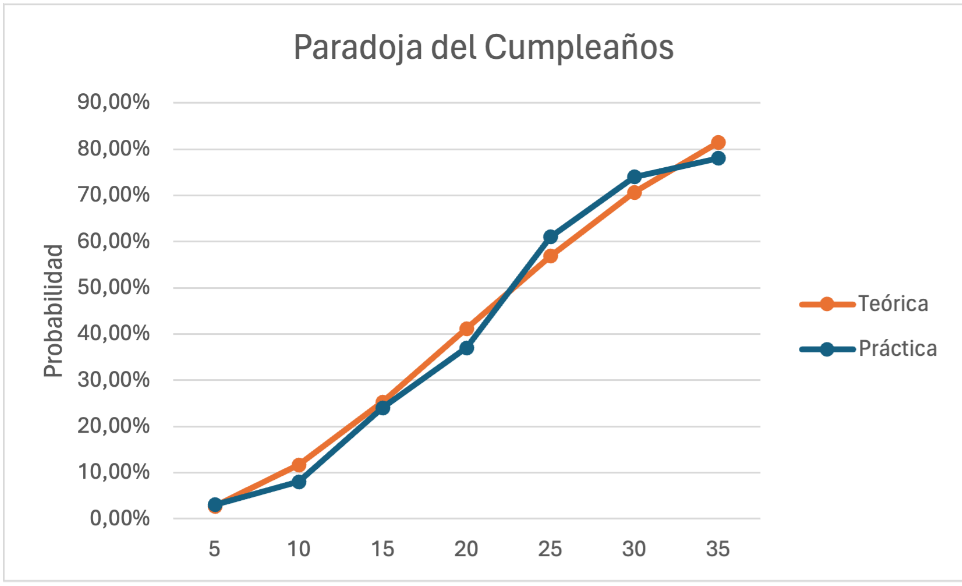
Aquí contrastaremos la probabilidad teórica para distintos valores de n en comparación a el testeo práctico para 100 intentos iguales para un mismo valor de n.



La room que se ve en la imagen es la sala número 100 recorrida para una sala de 23 personas

Número de personas (n)	Probabilidad Teórica	Probabilidad Práctica en 100 eventos
5	2.71 %	3 %
10	11.69 %	8 %
15	25.29 %	24 %
20	41.14 %	37 %
25	56.87 %	61 %
30	70.63 %	74 %
35	81.44 %	78 %

Cuadro 1: Comparación de probabilidades teóricas y prácticas



Eje X: Valores de n

Podemos ver que en general para distintos valores de n la probabilidad de teórica y la empírica se asemajan bastante para 100 intentos, en algunas ocasiones una es superior a la otra y viceversa, podemos ver que en general no existe una tendencia por que alguna de las dos sea consistenmente mayor o menor.

En general este fenómeno se puede explicar por la ley de los grande números, donde podemos ver que para 100 intentos ya nos estamos aproximando bastante bien a la probabilidad teórica, sin embargo si rehicieramos la prueba con 10.000 intentos para cada valor de n, estaríamos aun más cercanos al valor teórico.

Parte 2

c)

Cálculo teórico
Para la teória de esta probabilidad ocuparemos la Relación Gamma- Exponencial, donde T_i: tiempo entre cada evento y T_n: tiempo entre n eventos. Así podemos expresar a X: tiempo en se demoran responder a n alumnos de ingeniería $\sim Gamma(n, \mu)$ y Y: tiempo en se demoran responder a k alumnos de college $\sim Gamma(k, \tau)$. Entonces la probabilidad de que que 3 personas de Ingeniería reciban respuesta antes que 2 personas de College es equivalente a $P(Gamma(3, \mu) < Gamma(2, \tau))$

Cálculo empírico
Haciendo el cálculo en R podemos ver que conseguimos estas probabilidades para distintos valores de μ y τ

	τ = 0.1	τ = 0.6	τ = 1.1	τ = 1.6	τ = 2.1	τ = 2.6	τ = 3.1	τ = 3.6	τ = 4.1	τ = 4.6
μ = 0.1	0.3084	0.0109	0.0024	0.0009	0.0004	0.0003	0.0001	0.0000	0.0000	0.0000
μ = 0.6	0.9020	0.3187	0.1256	0.0669	0.0387	0.0242	0.0156	0.0100	0.0079	0.0049
μ = 1.1	0.9620	0.5524	0.3172	0.1908	0.1172	0.0785	0.0564	0.0445	0.0301	0.0259
μ = 1.6	0.9828	0.7078	0.4717	0.3191	0.2173	0.1612	0.1172	0.0894	0.0696	0.0564
μ = 2.1	0.9875	0.7815	0.5765	0.4171	0.3203	0.2343	0.1864	0.1473	0.1163	0.0974
μ = 2.6	0.9933	0.8414	0.6582	0.5035	0.3945	0.3106	0.2422	0.2003	0.1643	0.1346
μ = 3.1	0.9936	0.8831	0.7185	0.5753	0.4683	0.3778	0.3146	0.2508	0.2195	0.1794
μ = 3.6	0.9960	0.8994	0.7668	0.6384	0.5327	0.4414	0.3734	0.3189	0.2643	0.2283
μ = 4.1	0.9967	0.9207	0.8009	0.6809	0.5754	0.4905	0.4237	0.3648	0.3187	0.2667
μ = 4.6	0.9975	0.9319	0.8337	0.7256	0.6303	0.5438	0.4717	0.4028	0.3518	0.3067

```
mu <- 0.1
tau <- 0.1

integrand <- function(x) {
```

```
  dgamma(x, shape = 3, scale = mu) * pgamma(x, shape = 2, scale = tau)
}

prob <- integrate(integrand, lower = 0, upper = Inf)\$value
prob
```

d)

Cálculo teórico listings

Para calcular este ejercicio, podemos seguir con la misma dinámica del ejercicio anterior, donde ahora debemos calcular $P(Gamma(1, \mu) + Gamma(1, \tau) < Gamma(5, \tau))$ para distintos valores de μ y τ

Cálculo empírico

```
calculate_probability <- function(mu, tau) {
  mix <- 1 / (1 / mu + 1 / tau)

  integrand <- function(x) {
    dgamma(x, shape = 2, scale = mix) * pgamma(x, shape = 5, scale = tau)
  }

  prob <- integrate(integrand, lower = 0, upper = Inf)$value
  return(prob)
}

mu <- 0.1
tau <- 0.1
prob <- calculate_probability(mu, tau)
prob
```

en ese sentido, vemos que para estos valores, obtenemos como resultado 0.01783265.