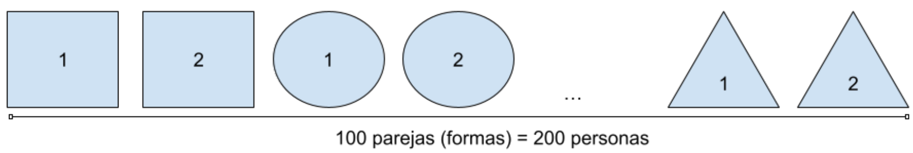
1. Problemas

1.1. (1 punto) *(10 minutos)* En un crucero transatlántico 100 parejas de recién casados han sido seleccionadas para formar un grupo de bailes de bachata. Si el grupo está compuesto por 14 personas, ¿Cuál es la probabilidad de que no haya parejas de casados en el grupo? Resuelve mediante el uso de combinatoria (NO se permite el uso de simulaciones).

Solution: Este problema es idéntico al problema 11 del tema de combinatoria (también resuelto en clase).

Para resolver el problema fácilmente, podemos imaginarnos a cada pareja como una forma geométrica (cuadrado está casado con cuadrado, triángulo con triángulo, etc.). Para distinguir a los miembros de cada pareja, podemos ponerles un identificador a cada cónyuge: cónyuge 1 y cónyuge 2.



Empezamos contando el número de casos totales, que es más sencillo: tenemos 200 personas totalmente distintas (equivalente a los 200 objetos que nos hemos imaginado) y para formar el grupo de bachata no nos importa el orden de elección:

$$\text{n}^{\underline{\text{o}}} \text{ casos totales} = \binom{200}{14}.$$

Para el número de casos favorables, lo mejor es proceder en dos pasos, comenzando con la restricción más importante del caso favorable: No podemos elegir a dos miembros de una pareja o, dicho de otra forma, los 14 objetos que usemos tienen que ser distintos.

- Paso 1: Elijo 14 formas de las 100 disponibles. Como no nos importa el orden, tenemos $\binom{100}{14}$. Un posible resultado de este paso se muestra en la siguiente figura.

- Paso 2: ¿Qué falta en la segunda figura con respecto a la primera? Para cada forma geométrica, necesitamos decidir a qué cónyuge incluimos (¡tenemos que elegir un número!). Como para cada objeto tenemos dos opciones:

$$\underline{2} \cdot \underline{2} \cdot \underline{2} \cdots \underline{2} = 2^{14}$$

Combinando los pasos 1 y 2, tenemos que

$$\text{n}^{\underline{\text{o}}} \text{ casos favorables } = \binom{100}{14} 2^{14}$$

:

Por lo tanto, la probabilidad queda:

$$p = \frac{\binom{100}{14} 2^{14}}{\binom{200}{14}} \approx 0,6136$$

Para calcular esa expresión

```
choose(100,14) * 2^14 / choose(200,14)
```

1.2. (2.5 puntos) *(20 minutos)* **[Entregar cumple.R]** A una fiesta de cumpleaños acuden 4 niños. Los organizadores de la fiesta llevan `n_caramelos`, y los reparten al azar entre todos los niños.

(a) (1.25 puntos) Escribe una función con prototipo `compute_prob(n_caramelos, n_sims)` que calcule, mediante SIMULACIONES, la probabilidad de que todos los niños reciban dos o más caramelos (`n_sims` hace referencia al número de simulaciones empleadas para el cálculo).

(b) (1.25 puntos) Usando la función del apartado anterior, ¿cuántos caramelos se deben llevar, como mínimo, para garantizar que todos los niños reciban dos o más caramelos con probabilidad superior a 0.95? Pista: usa un número de simulaciones bajo (p.ej.: 1000) para acotar los valores plausibles en una primera búsqueda. Incrementa luego el número de simulaciones para poder responder con precisión.

Solution: a)
En el ejercicio 3 de la práctica repartíamos `n_datablocks` en `n_servers`, aquí tenemos que repartir `n_caramelos` a 4 niños. La simulación comienza, por tanto, de forma análoga: para cada caramelo decidimos a qué niño se lo damos: `sample(n_ninos, n_caramelos, replace = TRUE)`.

A continuación, contamos usando `table` (usando el truco habitual de convertir a factor) y comprobamos si todos los niños han recibido dos o más caramelos (`all(table(reparto) >= 2)`.

Usando `replicate` y `mean` tenemos la solución completa.

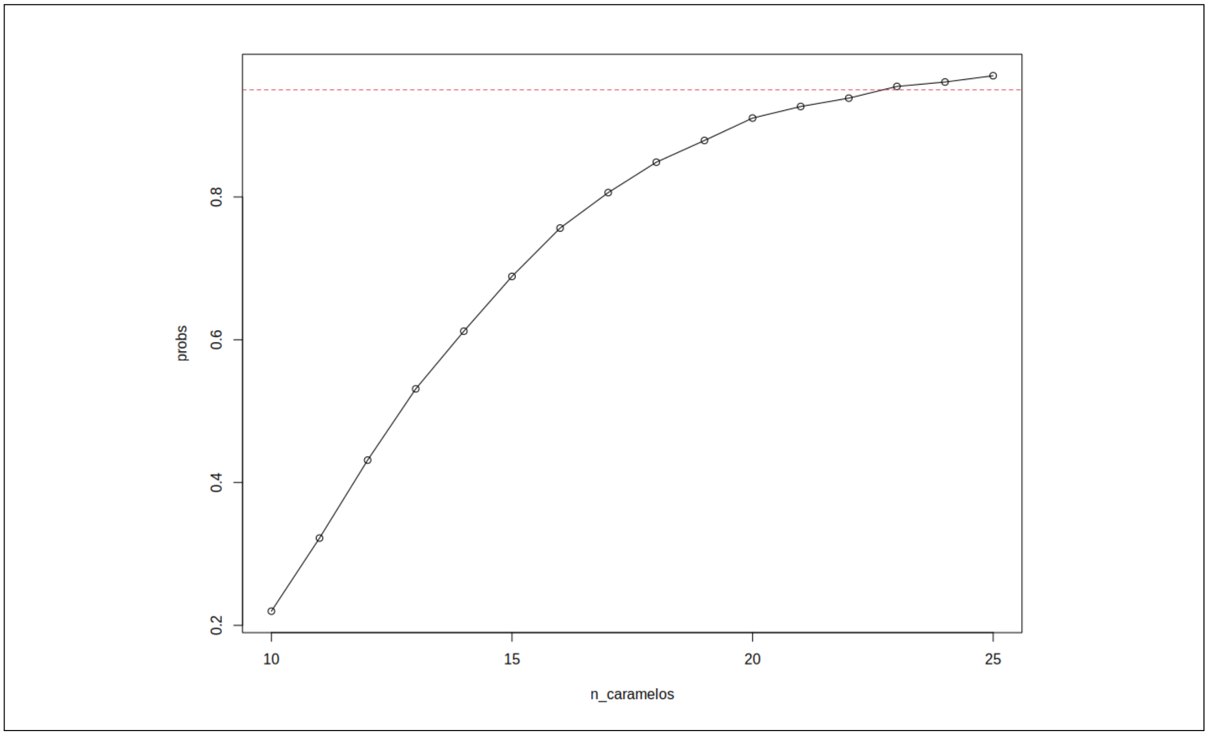
```r
n_caramelos = 10

compute_prob = function(n_caramelos, n_sims) {
  n_ninos = 4
  sims = replicate(n_sims, {
    reparto = sample(n_ninos, n_caramelos, replace=TRUE)
    reparto = factor(reparto, levels = 1:n_ninos)
    all(
      table(reparto) >= 2
    )
  })
  mean(sims)
}

compute_prob = Vectorize(compute_prob)
```

b)
Hay varias formas de resolver este apartado, pero la forma más visual es dibujar la probabilidad de que los niños reciban dos o más caramelos Vs. `n_caramelos`, y buscar qué punto cumple las condiciones del enunciado (probabilidad ≥ 0.95)

```r
n_caramelos = 10:25
probs = compute_prob(n_caramelos, n_sims = 10000)
plot(n_caramelos, probs, type = "o")
abline(h = 0.95, lty = 2, col = 2)
# Solución: 23 caramelos
```


probs
0.2
0.4
0.6
0.8
10
15
20
25
n_caramelos

1.3. (2.5 puntos) (25 minutos) Resuelve las siguientes preguntas relativas a distribuciones notables. Para cada apartado, selecciona distribuciones adecuadas para modelar las variables aleatorias involucradas y justifica brevemente tu elección.

(a) (1 punto) Cierta computadora se encarga del envío de mensajes militares desde un submarino. Debido a las condiciones extremas en las que opera el submarino, un mensaje no llega a su destino con probabilidad 0.7. Por otra parte, y para asegurarse que el contenido del mensaje es correcto, es necesario recibir 3 veces el mensaje para considerar una transmisión como exitosa. ¿Cuál es el número esperado de mensajes fallidos hasta conseguir una transmisión exitosa? ¿Y el número esperado de mensajes totales? NO está permitido el uso de simulaciones.

(b) (1.5 puntos) **[Entregar nodos.R]** Una red de telecomunicaciones tiene dos nodos, A y B, que manejan el tráfico de datos. El nodo A procesa el 30 % del tráfico, con un tiempo de procesamiento que sigue una distribución exponencial de media 5 milisegundos. El nodo B procesa el 70 % restante del tráfico y tiene un tiempo de procesamiento uniforme entre 1 y 7 milisegundos. Si se registra un procesamiento de menos de 3 milisegundos, ¿cuál es la probabilidad de que el tráfico haya sido procesado por el nodo A? Usa SIMULACIONES para resolver este apartado.

Solution: a)
Sea X: "nº de mensajes fallidos hasta conseguir 3 mensajes exitosos". Esto es la definición de una Binomial Negativa, por lo que:

$$X \sim \text{NegBinom}(3, 0.3).$$

Como el enunciado nos pregunta por "número **esperado** de mensajes fallidos..." debemos calcular $\mathbb{E}[X]$. Consultando el chuletario:

$$\mathbb{E}[X] = \frac{rq}{p} = \frac{3 \cdot 0.7}{0.3} = 7 \text{ mensajes fallidos.}$$

Dado que esperamos 7 mensajes fallidos hasta conseguir mandar 3 mensajes exitosos, en total esperamos mandar 10 mensajes.

b)
El inicio de este problema es análogo al de "Buffer overflow" de la práctica (en el que había días normales u ocupados y, en función del tipo de día, recibíamos más o menos mensajes). Las variables son:

- X: "nodo que realiza el procesamiento (0 significa que procesa A y 1 significa que procesa B)",
- Y: "tiempo de procesamiento en milisegundos"

y sus distribuciones:

$$\begin{aligned} &X \sim \text{Bernouilli}(0.7) = \text{Binom}(1, 0.7) \\ &Y \mid X = 0 \sim \text{Exp}(1/5) \\ &Y \mid X = 1 \sim \text{Unif}(1, 7) \end{aligned}$$

Para generar las simulaciones haremos...

```
n_sims = 5000
x = rbinom(n_sims, 1, 0.7)
y = ifelse(
    x == 0,
    rexp(n_sims, 1 / 5),
    runif(n_sims, 1, 7)
)
```

Finalmente, el enunciado nos pregunta por

$$P(X = 0 \mid Y < 3) = \frac{P(X = 0, Y < 3)}{P(Y < 3)} \text{ (fijaos en la pista de “Si se registra...”)}$$

y para calcularlo debemos seguir un procedimiento análogo al seguido en el ejercicio final de la práctica.

Todo el código completo sería:

```
n_sims = 5000
x = rbinom(n_sims, 1, 0.7)
y = ifelse(
  x == 0,
  rexp(n_sims, 1 / 5),
  runif(n_sims, 1, 7)
)

p_denom = mean(y < 3)
p_num = mean(
  (x == 0) & (y < 3)
)
print(
  p_num / p_denom
)
```

1.4. (3 puntos) *(30 minutos)* **[Entregar discreta.R]** La función de probabilidad conjunta de dos variables aleatorias X e Y se define como:

$$f(x,y) = \begin{cases} c(x^2+2y^2) & \text{si } xy = 12,\ x,y \in \mathbb{Z} \\ 0 & \text{en otro caso} \end{cases}$$

Pista: Ten en cuenta que $\mathbb{Z} = \{\dots, -3, -2, -1, 0, +1, +2, +3, \dots\}$.

(a) (0.5 puntos) Calcula la constante c. Si no eres capaz de resolver este apartado, usa $c = 1$ en los siguientes apartados.

(b) (0.5 puntos) Calcula $P(X < Y + 5)$.

(c) (1 punto) Calcula las funciones de probabilidad marginales de X e Y y represéntalas gráficamente.

(d) (1 punto) Calcula $\mathbb{E}[X+Y]$ y $\mathbb{E}[X \cdot Y]$.

Solution: Tal y como sugiere la pista del enunciado ("Ten en cuenta que ...") y el nombre del ejercicio (**discreta.R**), la distribución es discreta. ¡No hay que hacer integrales!

```r
### (a)
# Función conjunta f no normalizada (sin la constante)
fxy_unnormalized = function(x, y) {
  ifelse(x * y == 12, x ^ 2 + 2 * y ^ 2 , 0)
}
# Tabla de probabilidades
probs_unnormalized = outer(-12:12,-12:12,fxy_unnormalized)

# De c * sum(f_unnormalized) = 1 obtenemos:
c_value = 1 / sum(probs_unnormalized)

fxy = function(x,y) {
  c_value * fxy_unnormalized(x, y)
}

### (b)
# Generamos la matriz de probabilidades
probs = outer(-12:12,-12:12, fxy)

# La idea es sumar aquellas probabilidades que cumplan
# la condición del enunciado (X < Y + 5)
is_in_domain = function(x, y) {x < y + 5}
# P(X < Y + 5)
sum(
  probs[outer(-12:12, -12:12, is_in_domain)]
)

# Una forma alternativa es defininir una función conjunta
# restringida al dominio X < Y + 5
fxy_restringida = function(x,y){
  ifelse((x*y == 12) & (x < y + 5), c_value * (x^2 + 2 * y^2) ,0)
}
# P(X < Y + 5)
sum(outer(-12:12,-12:12, fxy_restringida))
```

```r
### (c)
# Función de densidad marginal de X y plot
px = rowSums(probs)
plot(-12:12, px, type = "h")

# Función de densidad marginal de Y y plot
py = colSums(probs)
plot(-12:12, py,type = "h")

### (d)
# para resolver tanto E[X + Y] como E[X * Y] por el mismo
# procedimiento usaremos la ley del estadístico inconsciente.
# E[X + Y]
summation_term = function(x,y){
  (x + y) * fxy(x, y)
}
sum(outer(-12:12,-12:12, summation_term))

## E[XY]
summation_term = function(x,y){
  (x * y) * fxy(x, y)
}
sum(outer(-12:12, -12:12, summation_term))
```

1.5. (1 punto) *(10 minutos)* De una variable aleatoria X se sabe que su esperanza es 6 y su varianza es 12. Escribe su función de probabilidad sabiendo que X es una binomial negativa.

Solution: Basta con utilizar las fórmulas de la esperanza y varianza que hay en el chuletario y despejar.

$$\begin{cases} r\dfrac{q}{p} = 6 \\ r\dfrac{q}{p^2} = 12 \end{cases}$$

Llevamos $6 = r\frac{q}{p}$ a la segunda ecuación:

$$r\frac{q}{p^2} = r\frac{q}{p}\frac{1}{p} = 6\frac{1}{p} = 12$$

de donde $p = 1/2$ (y, por tanto, $q = 1/2$). Usando esta información en la primera ecuación se obtiene que $r = 6$.

Dado que $(r, p) = (6, \frac{1}{2})$, la función de densidad queda (ver fórmula del chuletario):

$$f(x) = \begin{cases} \binom{x+5}{5}\frac{1}{2^{x+5}} & \text{si } x = 0, 1, 2, \dots \\ 0 & \text{en otro caso} \end{cases}$$