## Handout 2

1. Assume you have a regular coin ($P(\text{heads}) = P(\text{tails}) = 0.5$).
   (a) What is the probability that you get 6 or more heads in 10 throws?
   (b) What is the probability that you get between 4 and 7 heads in 10 throws?
2. Using tables, compute the following probabilities:
   (a) $P(X \leq 3)$ when $X \sim b(p = 0.3, n = 15)$.
   (b) $P(1 < X \leq 3)$ when $X \sim b(p = 0.3, n = 15)$.
   (c) $P(X \leq 5)$ when $X \sim b(p = 0.7, n = 20)$.
   (d) $P(3 \leq X < 5)$ when $X \sim b(p = 0.7, n = 20)$.
3. If the probability of a couple having a child of either sex were the same, what would be the probability that in families with five children there were exactly two girls? At most two girls?
4. Use the moment generating (or characteristic) function to prove or disprove that the sum of two independent binomial random variables with respective distributions $b(p_1, n_1)$ and $b(p_2, n_2)$, $p_1 \neq p_2$, is $b(p_1 + p_2, n_1 + n_2)$.
5. Let $Z = X_1 + X_2$ with $X_1 \sim b(p_1 = 0.5, n_1 = 10)$ and $X_2 \sim b(p_2 = 0.5, n_2 = 5)$, independent of each other. Find $P(Z \leq 7)$.
6. Let $Z = X_1 + X_2$ with $X_1 \sim b(p_1 = 0.5, n_1 = 2)$ and $X_2 \sim b(p_2 = 0.3, n_2 = 2)$, independent of each other. Find $P(Z \leq 1)$, both using and not using moment generating functions.
7. A batch of 10000 parts is sampled to check the quality. According to the seller, a fraction of no more than 10% are defectives. What is the probability, if indeed there are 10% of defectives, that in a random sample of $n = 20$ we find more than three? Is the distribution of defectives in the sample of $n = 20$ really binomial? Why or why not?
8. A company operates a fleet of buses on long distance hauls. It has been observed that on average one out of five passengers asks for a given newspaper during the trip. How many copies does the attendant of a bus with 20 passengers need to purchase so that demand will be met with probability greater than 95%?

## Pautas docentes

1. En el ejercicio 2 se puede, al mismo tiempo que mostramos el uso de las tablas (suelo tenerlas proyectadas), hacer lo mismo con R (manteniendo abierta una sesión en Rstudio y alternándola en la pantalla con las tablas).

   El cálculo en R es fácil:

   ```
   > pbinom(q=3, size=15, prob=0.3)

   [1] 0.2968679

   > pbinom(q=3, size=15, prob=0.3) - pbinom(q=1, size=15, prob=0.3)

   [1] 0.2616003

   > pbinom(q=5, size=20, prob=0.7)

   [1] 4.294002e-05

   > pbinom(q=4, size=20, prob=0.7) - pbinom(q=2, size=20, prob=0.7)

   [1] 5.512522e-06
   ```

   De pasada podemos introducir el uso de `dbinom` y comparar con el último resultado:

   ```
   > dbinom(x=3, size=20, prob=0.7) + dbinom(x=4, size=20, prob=0.7)

   [1] 5.512522e-06
   ```

   o formas alternativas de hacer lo mismo:

   ```
   > sum( dbinom(3:4, size=20, prob=0.7) )

   [1] 5.512522e-06
   ```

2. Resaltar que la nomenclatura es fácil de recordar: para cada distribución, cuatro funciones comenzando por `d,p,q,r` que proporcionan respectivamente densidad (o probabilidad), distribución, función cuantil (ó $F_X^{-1}(q)$) y generador de números aleatorios. Por ejemplo, `dbinom`, `pbinom`, `qbinom` y `rbinom`.

3. Resaltar la diferencia entre los ejercicios 5 y 6: si el parámetro $p$ de las binomiales que se suman no es el mismo, el resultado no es binomial.

4. El ejercicio 7 da pie a introducir la distribución hipergeométrica. La probabilidad exacta del suceso que se dice es

$$1 - \left[ \frac{\binom{9000}{20}\binom{1000}{0}}{\binom{10000}{20}} + \frac{\binom{9000}{19}\binom{1000}{1}}{\binom{10000}{20}} + \frac{\binom{9000}{18}\binom{1000}{2}}{\binom{10000}{20}} + \frac{\binom{9000}{17}\binom{1000}{3}}{\binom{10000}{20}} \right]$$

Llamar su atención sobre:

(a) Si una probabilidad es larga de obtener, podemos (como hemos hecho aquí) obtener la del suceso complementario y restar de 1.

(b) Incluso utilizando este atajo, la probabilidad deseada es larga de obtener, y la aproximación binomial es muy buena con muestras tan pequeñas en relación al tamaño de la población.

Formas alternativas de calcular la probabilidad exacta en R:

```
> 1 - phyper(3, m=1000, n=9000, k=20)

[1] 0.1327555

> sum( dhyper(4:20, m=1000, n=9000, k=20) )

[1] 0.1327555
```

Aproximación binomial:

```
> 1 - pbinom(3, size=20, prob=0.10)

[1] 0.1329533

> sum( dbinom(4:20, size=20, prob=0.10) )

[1] 0.1329533
```

5. El problema 8 es deliberadamente ambiguo y busca que ejerzan su criterio acerca de qué supuestos son razonables. En ningún momento se dice que las peticiones de los pasajeros sean independientes: si creen que lo son, la distribución binomial es utilizable. En otro caso (efecto emulación: "¿Hay 'El Correo'? Tráigame uno a mi también") la distribución podría ser muy alejada de la binomial.

Hay que calcular el primer número de periódicos que deja por debajo la demanda con probabilidad 0.95. Fácil con tablas o mediante:

```
> pbinom(0:20,size=20, prob=0.20)

 [1] 0.01152922 0.06917529 0.20608472 0.41144886 0.62964826 0.80420779
 [7] 0.91330749 0.96785734 0.99001821 0.99740517 0.99943659 0.99989827
[13] 0.99998484 0.99999815 0.99999982 0.99999999 1.00000000 1.00000000
[19] 1.00000000 1.00000000 1.00000000
```