Profesor: Héctor Bahamonde, PhD.

e:hector.bahamonde@uoh.cl

w:www.hectorbahamonde.com

Curso: MLE.

TA: Gonzalo Barría.

I. Probabilidad e Incertidumbre

Existen tres axiomas del modelo de probabilidad. Los axiomas no son ni verdaderos ni falsos, "simplemente son". Desde estos tres axiomas, se puede derivar todo el resto de la teoría de probabilidad.

1. Para cualquier evento z, $Pr(z) \geq 0$.
2. $Pr(\mathcal{C}) = 1$.
3. Si todos los eventos K son mutuamente independientes, $Pr(z_1 \cup z_2 \cup z_3 \cup \ldots z_K) = Pr(z_1) + Pr(z_2) + Pr(z_3) \ldots Pr(z_K)$

II. Distribuciones de Probabilidades Univariadas

En MLE, es fundamental pensar en el "underlying process generating [the data]". *Qué tipo de proceso está generando los datos?* Por ej., que una moneda de cara (1) o sello (0) no representa una distribucion contiunua. Se trata de otro "data generating process" y se podría explicar con la siguiente notación:

$$
\begin{aligned}
Pr(Y_i = 1) &= \pi \\
Pr(Y_i = 0) &= 1 - \pi
\end{aligned} \tag{1}
$$

donde π es el parámetro (lo que estimamos) de la distribución (el montón de 1's y 0's). Nuestro parametro en modelos lineales OLS era β. Ahora en MLE es π. El valor esperado de este evento z (cuando la moneda no está cargada) es 0.5, o más formalmente, $\pi = 0.5$. Por qué?

Este proceso se llama "Bernoulli".

I. Bernoulli

Bernoulli es un tipo especial de distribución "binomial" (de dos resultados). En general, siempre son 1's o 0's (pero puede ser cualquier otro numero—al final del día, es una distribución "categorica"). Ejemplos?

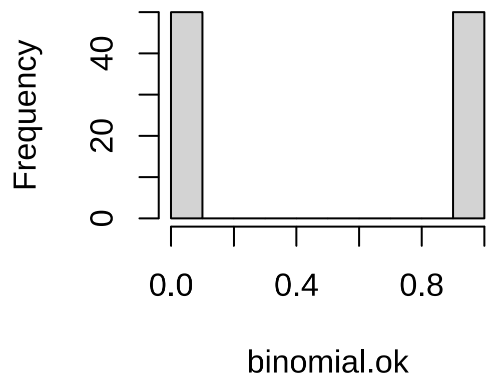
Lo importante es que:

1. La probabilidad de cada evento es mayor a 0.
2. Los posibles resultados son exhaustivos (?). O lo que es lo mismo, $Pr(Y = 1|Y = 0) = 0$.

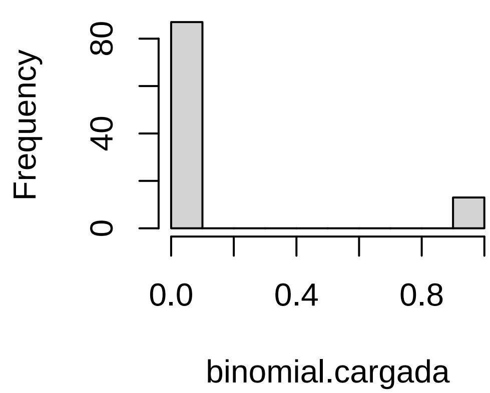
Más formalmente, la **fórmula de la distribución Bernoulli** es la siguiente:

$$f(y_i|\pi) = \begin{Bmatrix} \pi^{y_i}(1-\pi)^{1-y_i} & \text{for } y_i = 0, 1, \\ 0 & \text{otherwise} \end{Bmatrix} \tag{2}$$

```
# Normal
p_load(extraDistr)
binomial.ok =  rbern(100, # N
                     prob = 0.5 # probabilidad de obtener 1
                     )
hist(binomial.ok, main = "")
```




```r
# Cargada
binomial.cargada =  rbern(100, # N
                     prob = 0.1 # probabilidad de obtener 1
                     )
hist(binomial.cargada, main = "")
```



II. Binomial

Proceso Bernoulli, pero donde sólo se observa la sumatoria. Por ejemplo, si preguntas si haz votado en las ultimas tres elecciones, seria algo asi:

- Votaste en la eleccion 1: Sí (1) o No (0)? (proceso Bernoulli, pero **no observas esto**).
- Votaste en la eleccion 2: Sí (1) o No (0)? (proceso Bernoulli, pero **no observas esto**).
- Votaste en la eleccion 3: Sí (1) o No (0)? (proceso Bernoulli, pero **no observas esto**).
- **Sumatoria**: "2" (**sí observas esto**).

Es importante tener en cuenta que el supuesto es que las tres decisiones son **independientes**. Estamos dispuestos a creer en esto? Es una conveniencia. Si es así, podemos modelar este proceso usando el mismo π para todo el proceso binomial (i.e. cada uno de los procesos Bernoulli).

$$f(y_i|\pi) = \left\{ \begin{array}{cc} \frac{N!}{y_i!(N-y_i)} \pi^{y_i}(1-\pi)^{N-y_i} & \text{for } y_i = 0, 1, \ldots, n, \\ 0 & \text{otherwise} \end{array} \right\} \tag{3}$$

```
# 1 Eleccion (observas el total por cada individuo)
rbbinom(20, # N
size = 1 # "Trials" o pruebas
)

##  [1] 1 0 1 0 0 0 0 0 0 1 1 0 0 1 1 1 1 0 0 0

# 10 Elecciones (solo observas el total por cada individuo)
rbbinom(20,#N
size = 10 # "Trials" o pruebas
)

##  [1]  4  2  7 10  0  4  9  2  0  8  8  0  0  1 10  1  4  5  4  0
```

Aqui usamos el comando `rbbinom` que es para *beta-binomial* (no binomial). Esta distribución no asume que cada elección para cada individuo es independiente.

III. Poisson

Este proceso se refiere a cuentas positivas, discretas, y sin tope. Por ejemplo, la cantidad de veces que te haz comprado bebidas en tu vida. Si te pones a contar esas veces, harás una distribución Poisson (no puedes comprar negativas bebidas, no puedes en teoría haber comprado una fracción de vez, y no existe un tope al número que puedas comprar). **Tiene los siguientes supuestos**:

1. Comienza en cero.
2. Se asume que la ocurrencia de un evento es independiente de la ocurrencia de los otros eventos.
3. Sólo un evento ocurre a la vez.
4. Los periodos de observación tienen el mismo largo.
5. **Promedio y varianza son iguales**.

Ejemplos?

$$f(y_i|\lambda) = \begin{Bmatrix} \frac{\epsilon^{-\lambda}\lambda^{y_i}}{y_i!} & \text{for } \lambda > 0 \text{ and } y_i = 0, 1, \ldots, \\ 0 & \text{otherwise} \end{Bmatrix} \tag{4}$$

donde λ es el evento, y donde $\epsilon^{-\lambda} = Pr(y_i = 0)$, asumiendo que todos los t tienen el mismo largo.

```
rtpois(20, # N
       5 # lambda (que es el promedio en una dist Poisson)
       )

##  [1] 6 2 4 5 7 4 5 7 1 6 8 1 6 7 3 5 1 4 7 5
```

Aquí ocupamos `rtpois` que es para *random truncated Poisson* (es truncada porque trunca los valores—desde menos infinito hasta infinito positivo—pero sigue siendo una distribución Poisson).

IV. Negative Binomial

La distribución Poisson asume que el radio de ocurrencia del evento λ es constante ("promedio y varianza son iguales"). Es decir, siempre tienes en expectativa, más o menos los mismos valores (o ellos no están dispersos). Poisson también asume que observas todos los periodos. Qué haces cuando estos dos supuestos no se cumplen? Usas Negative-Binomial. Ejemplos?

$$f_{nb}(y_i|\lambda,\sigma^2) = \frac{\Gamma(\frac{\lambda}{\sigma^2-1}+y_i)}{y_i!\Gamma(\frac{\lambda}{\sigma^2})}(\frac{\sigma^2-1}{\sigma^2})^{y_i}(\sigma^2)^{\frac{-\lambda}{\sigma^2-1}} \tag{5}$$

donde Γ es la función Gamma donde $\Gamma(n) = (n-1)!$, y donde λ sigue siendo la ocurrencia de un evento.

```
rbnbinom(20, # N
1) # "Trials" o pruebas.

##  [1]  4  0  3  0  4  0  1  5  1  0 38  4  0  2  0  0  0  0  7  9
```

Fíjate como la varianza no es constante (hay números más grandes, como el "38").