

Enfoque Bayesiano de la estadística

Mario Carranza

Comenzamos con una pregunta

- Dentro de la biología, ¿existen escuelas de pensamiento que estén enfrentadas o tengan controversias?

Enfoques en la estadística

- Fisher- Verosimilitud (intervalos de verosimilitud, p-valor)
- Frecuentista- Neyman/ Pearson (intervalos de confianza, error tipo I, error tipo II)
- Bayesiano (Intervalos de credibilidad, factor de Bayes)

Personajes relevantes



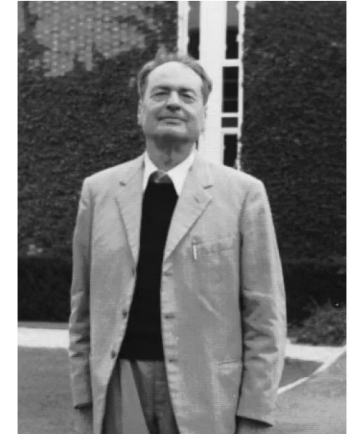
Thomas Bayes
(1701 –1761)



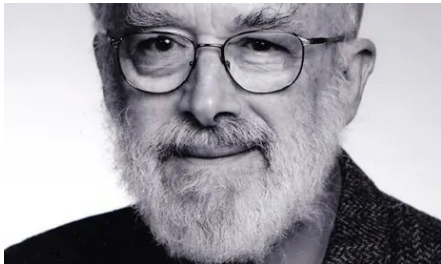
Simon Laplace
(1701 –1761)



Frank Ramsey
(1903-1930)



Bruno de Finetti
(1906-1985)



Dennis Lindley
(1923-2013)



Leonard J. Savage
(1917-1971)

El demonio de Laplace



Suponiendo el determinismo de Laplace

La ubicación y momento de cada partícula en el universo.

Las leyes de mecánica clásica

Todos sus valores pasados o futuros podrían calcularse.

Tres interpretaciones de la probabilidad

- Clásica: Resultados favorables/ Resultados posibles
- Frecuentista: límite de la razón de frecuencias.
- Subjetivista: Representación de la incertidumbre del tomador de decisiones

Otros casos de información previa


- Formula de Cochran para tamaños de muestra

$$n = \frac{n_0}{1 + \frac{n_0}{N}}$$

$$n_0 = \frac{(Z_{\alpha/2})^2 S^2}{\epsilon^2}$$

$$n_0 = \frac{Z^2 p(1-p)}{e^2},$$

Teoría de la decisión con incertidumbre

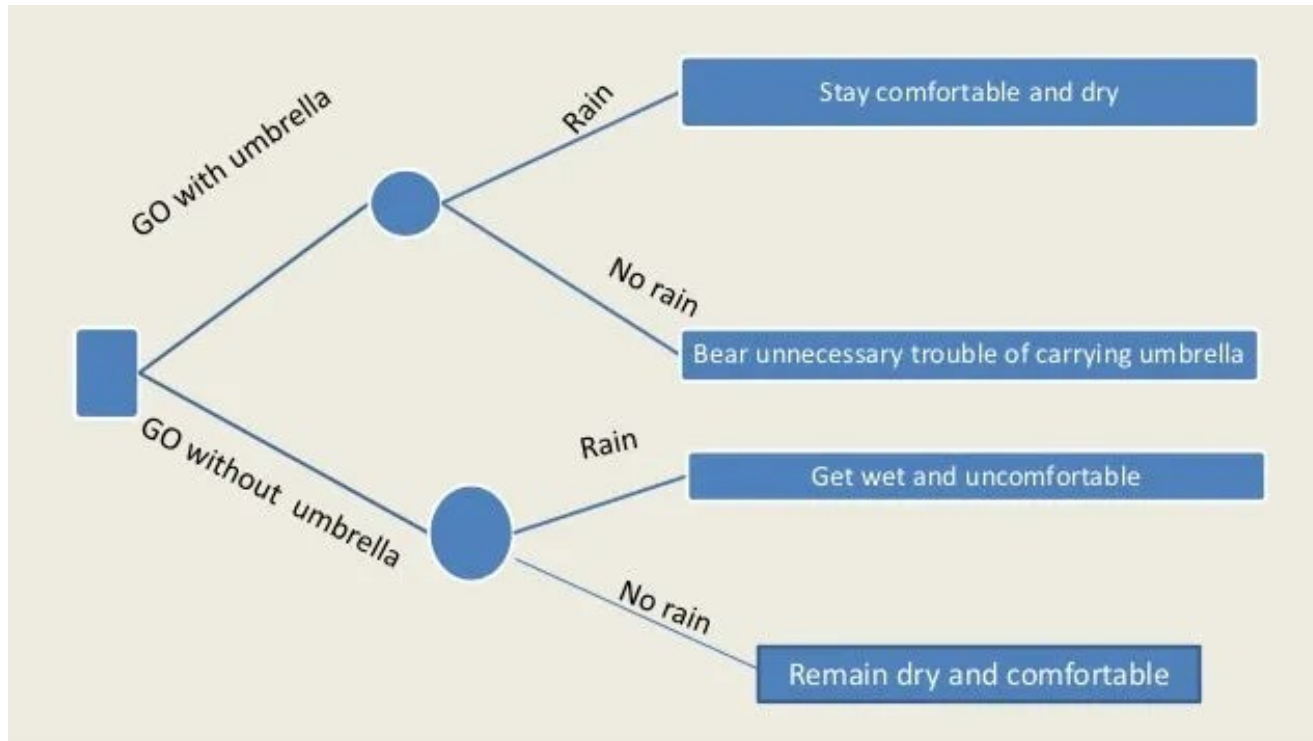


Espacio
de
decisiones

Espacio
de
estados
de la
naturaleza

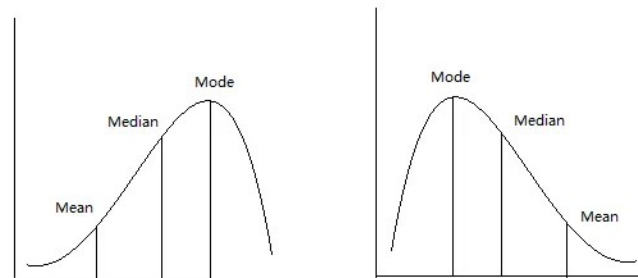
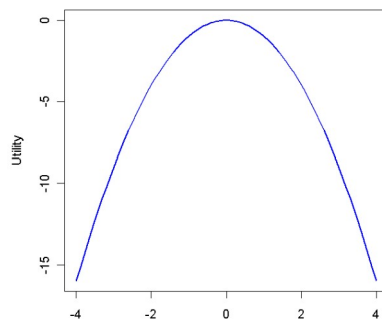
Espacio
de
consecuencias

Ejemplo de problema de decisión del paraguas

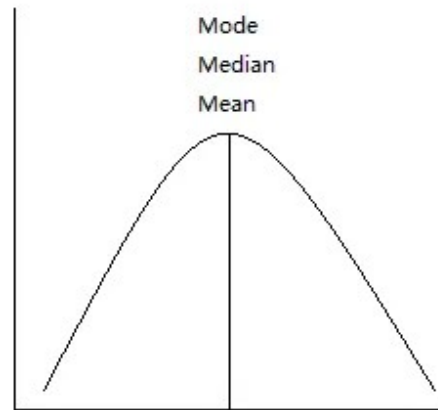


Inferencia como problema de decisión

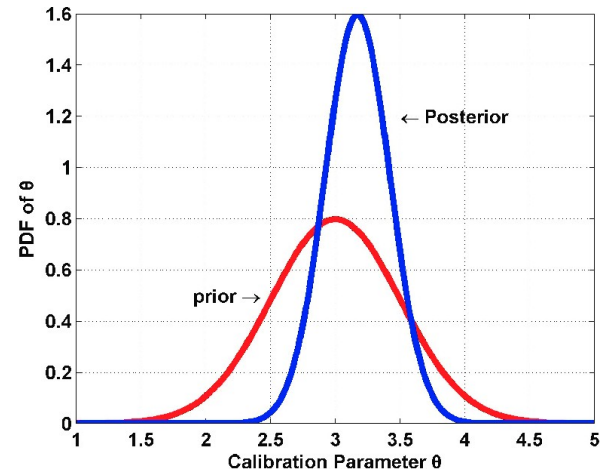
Quadratic Distribution Utility Function



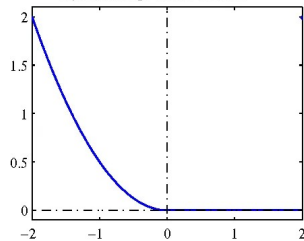
Left skew



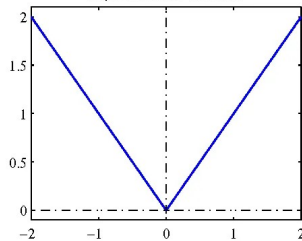
Normal Distribution



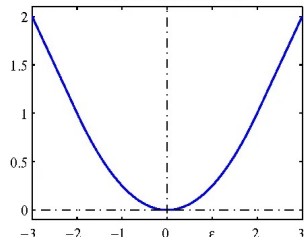
Squared Hinge Loss Function



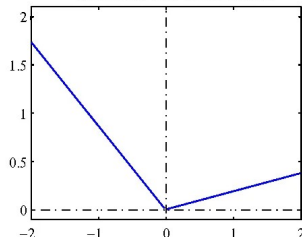
Laplacian Loss Function



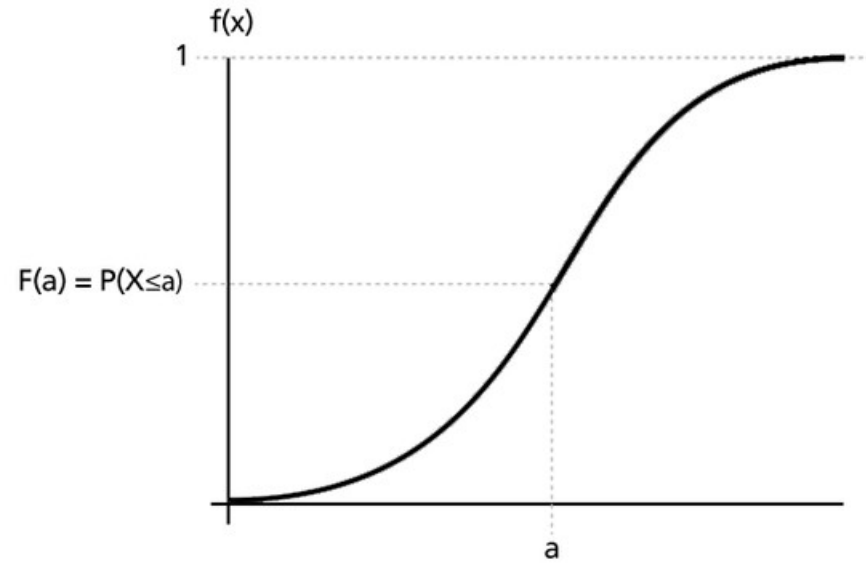
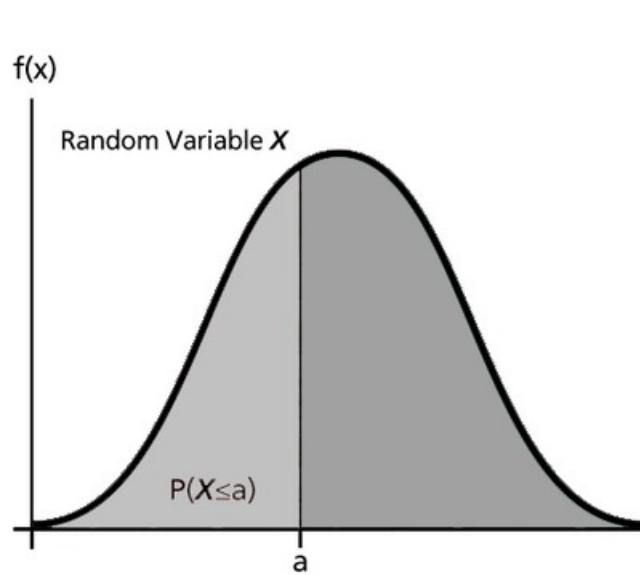
Huber's Loss Function



Quantile ($q=0.9$) Loss Function



Distingamos la densidad de la acumulada



Problemas estadísticos

- Estimación puntual: Dar el mejor valor que aproxime un parámetro
- Estimación por intervalos: Dar una región (usualmente entre dos valores) que contenga al verdadero valor del parámetro.
- Prueba de hipótesis: Contrastar la plausibilidad de una afirmación.

¿Cómo calculamos probabilidades?

- $X = \{X_1, \dots, X_n\}$ con $X_i \sim P(A; \Theta)$: datos
- $P(A; \Theta)$ Modelo
- Vemos que si $A=(a,b)$, $P(A; \Theta) = F(b; \Theta) - F(a; \Theta)$
- Hasta ahora ni pinta la $f(x; \Theta)$

En resumen

Modelo $P(X;\Theta)$

Datos
 X

Inferencia/ Inducción (Estadística)

Parámetros
 Θ

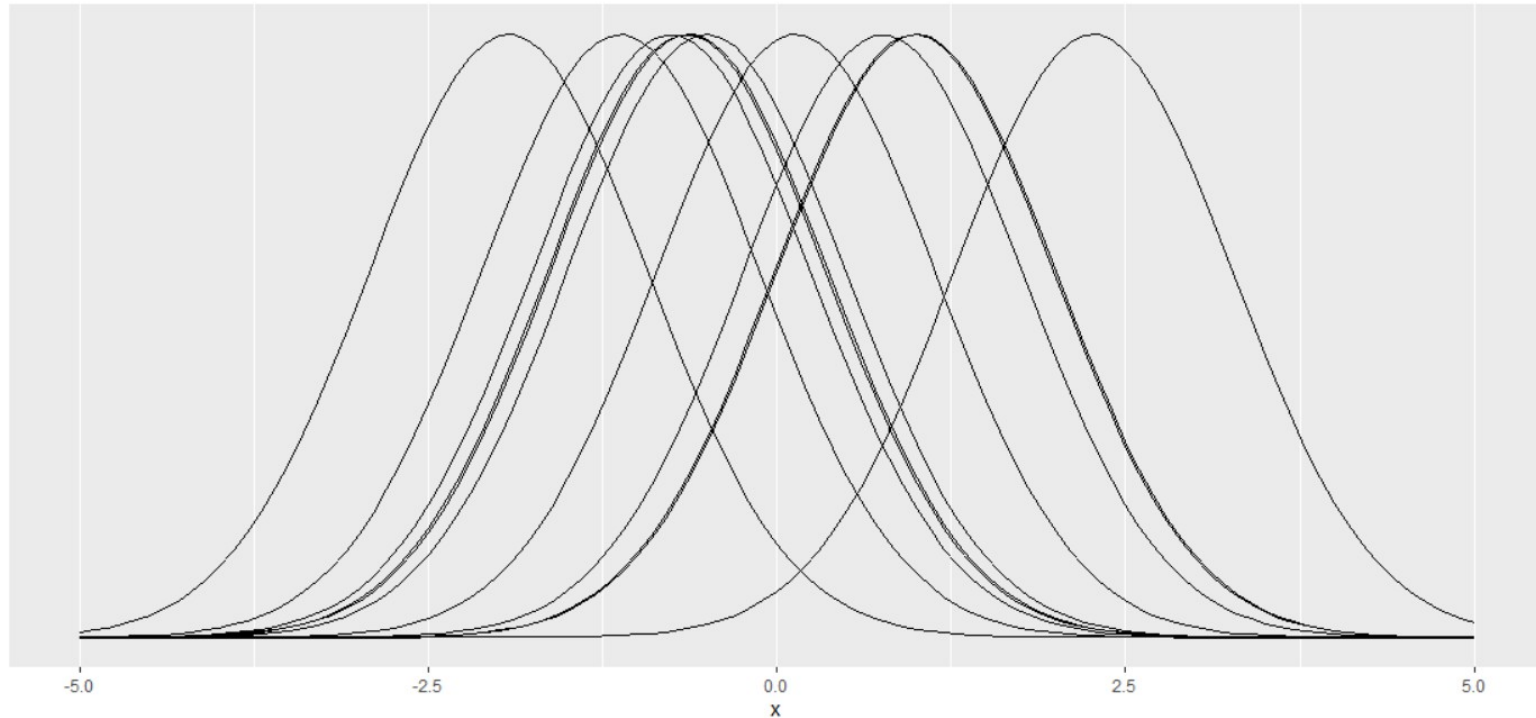
Pronóstico/ Predicción (Probabilidad)

Probabilidad $[0,1]$

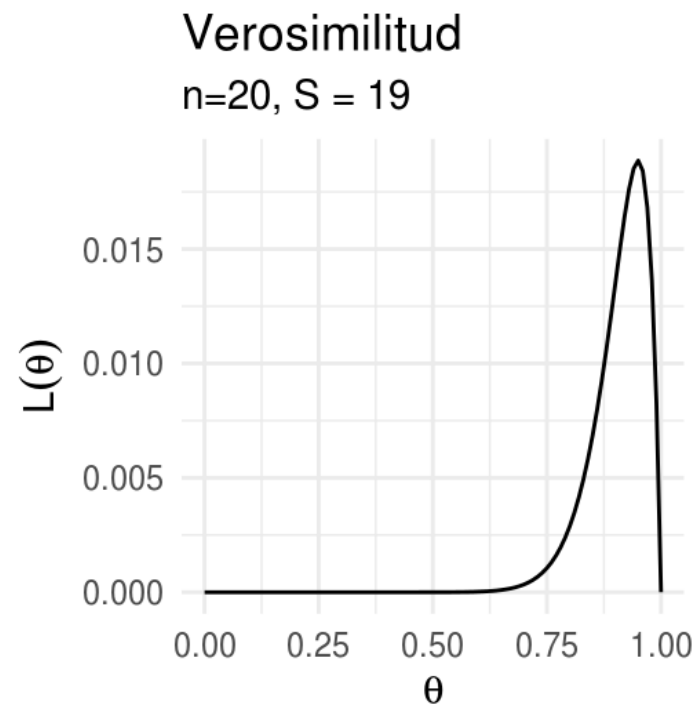
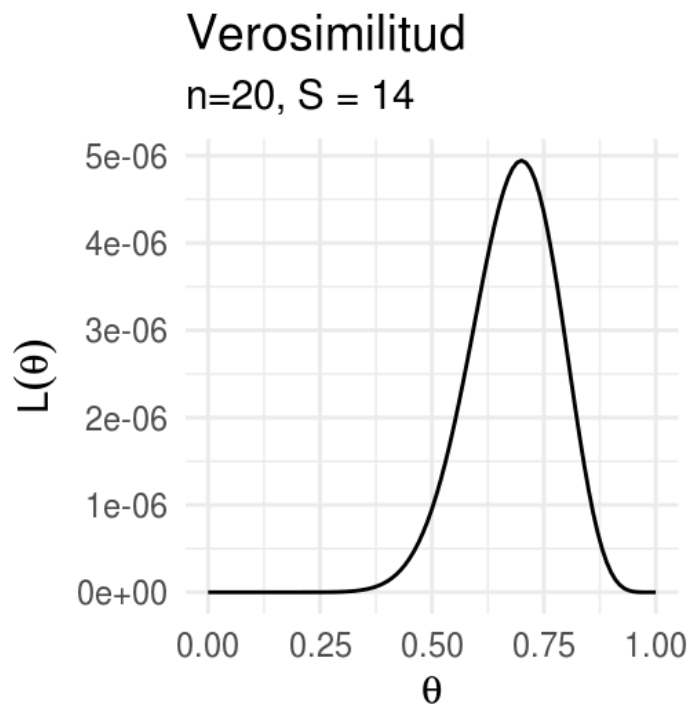
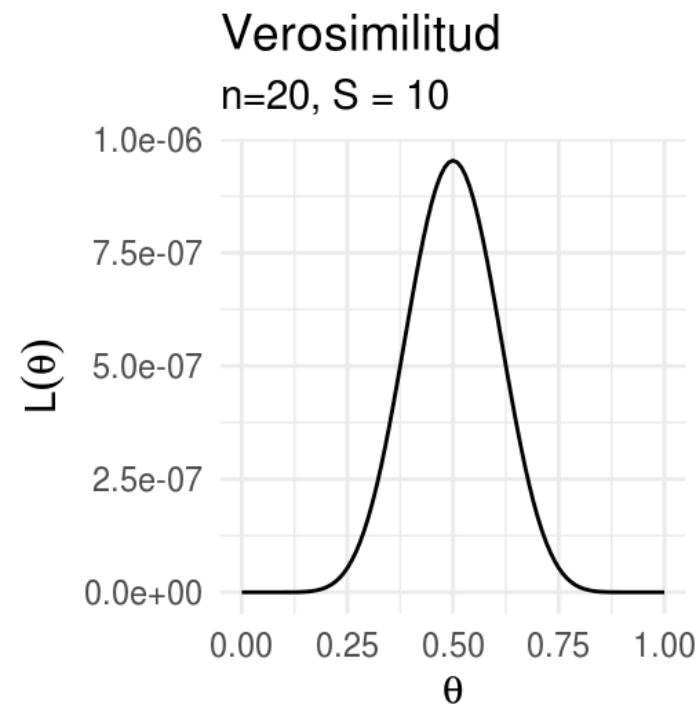
Resulta que la densidad es muy útil para la inferencia

- Función de Θ , dados los datos x_1, \dots, x_n
- Construimos una función multiplicando las densidades evaluadas en cada uno de los datos. (Si la muestra es independiente)
- $L(\Theta; X) = f(x_1; \Theta) * \dots * f(x_n; \Theta)$

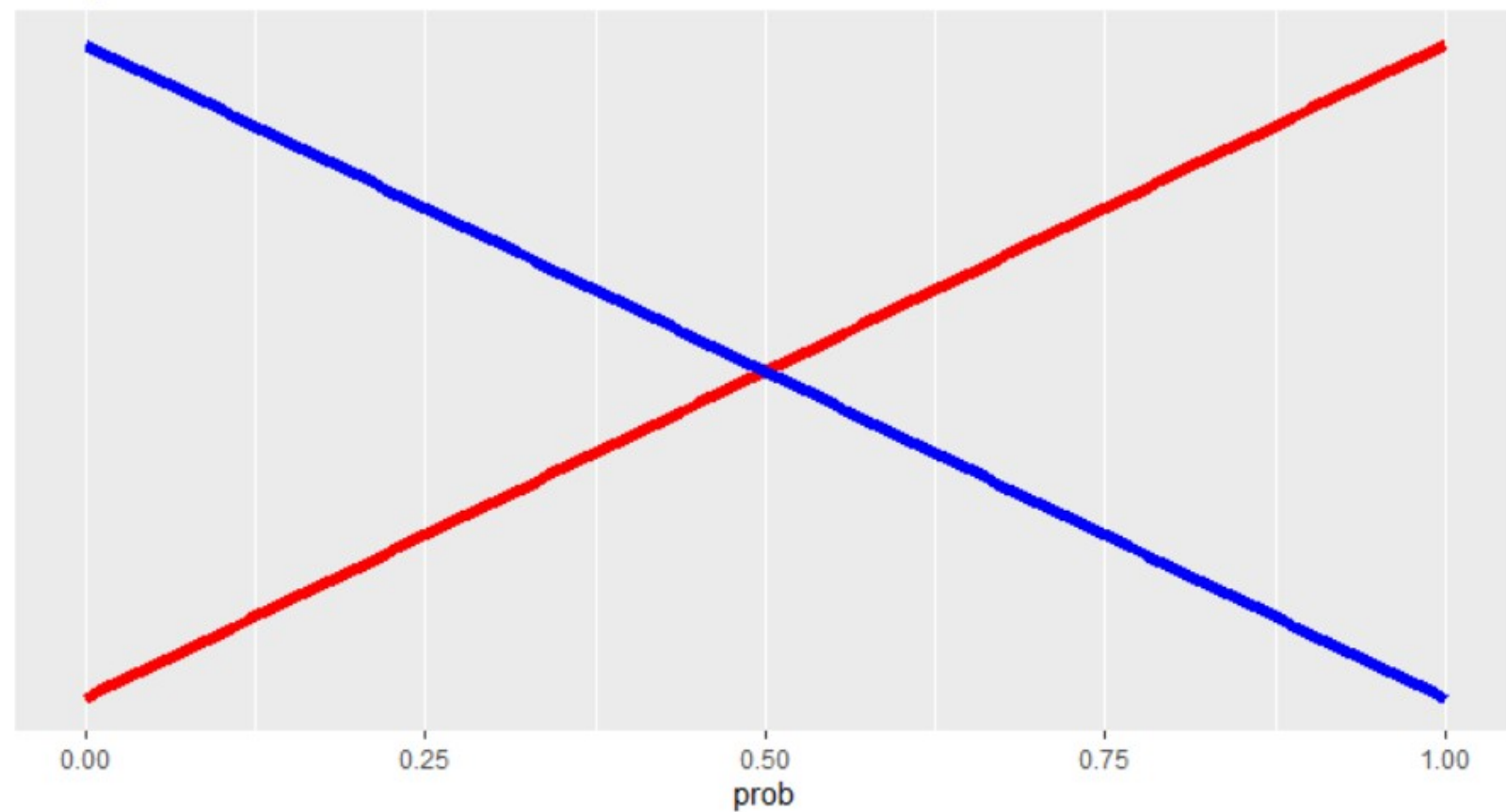
Densidades como función del parámetro



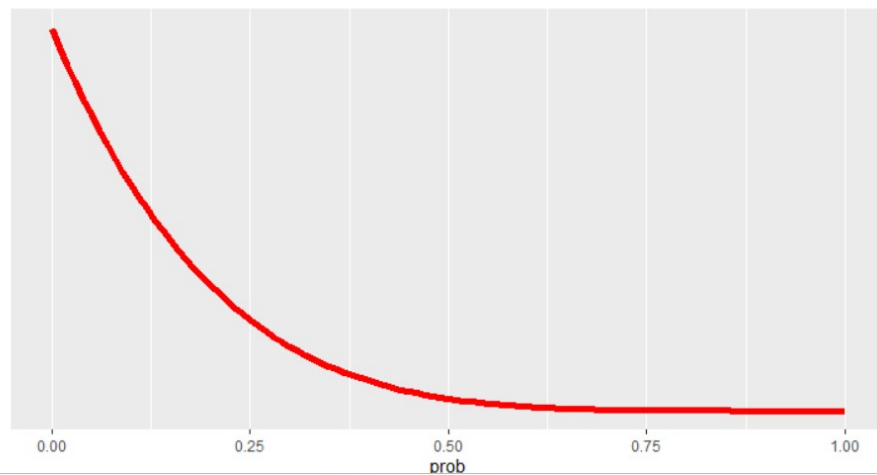
Esta función es muy útil en estimación



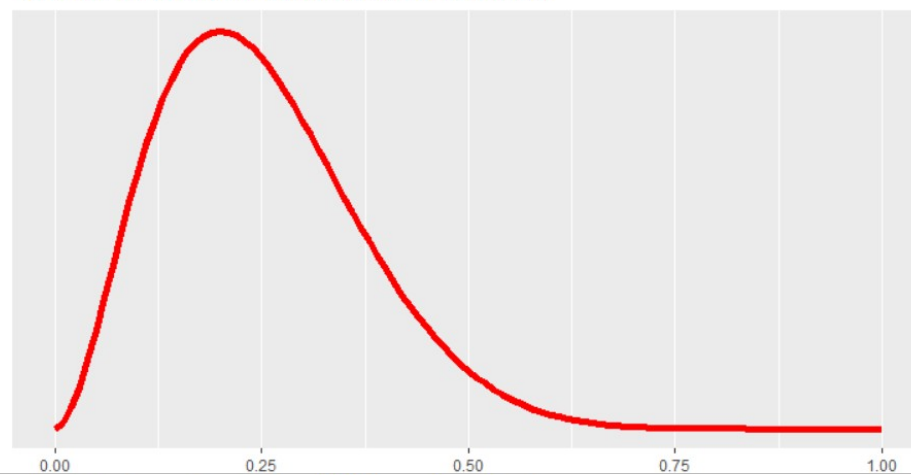
Componentes de la verosimilitud caso Volados o Prevalencia



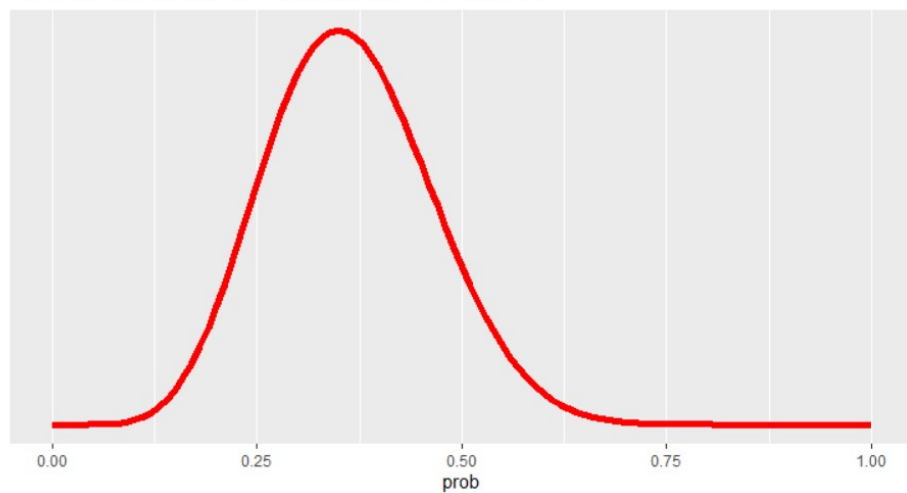
Tamaño de muestra 5 Caso Volados o Prevalencia



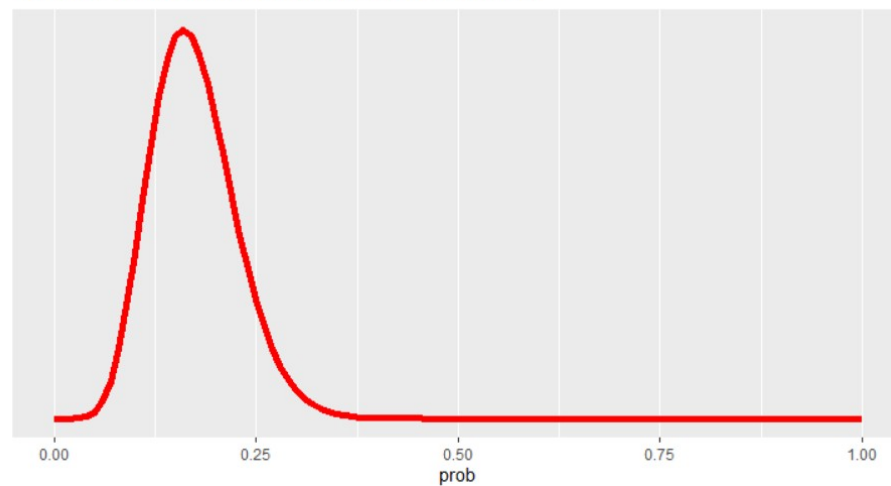
Tamaño de muestra 10 Caso Volados o Prevalencia



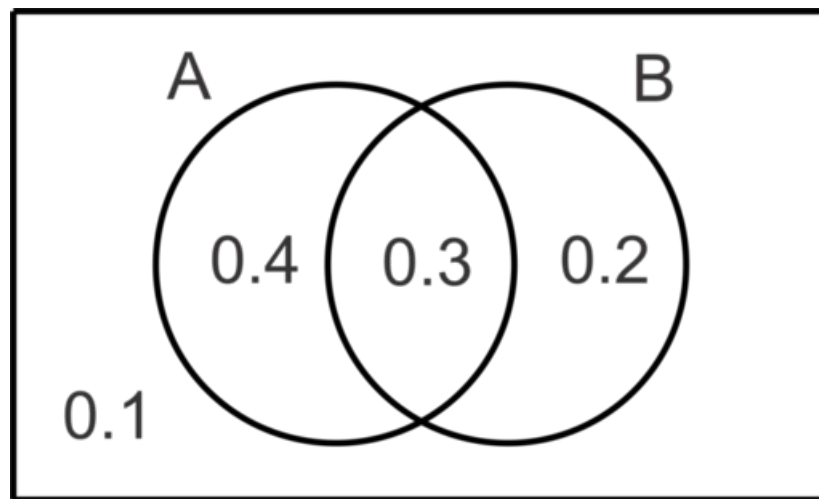
Tamaño de muestra 20 Caso Volados o Prevalencia



Tamaño de muestra 50 Caso Volados o Prevalencia

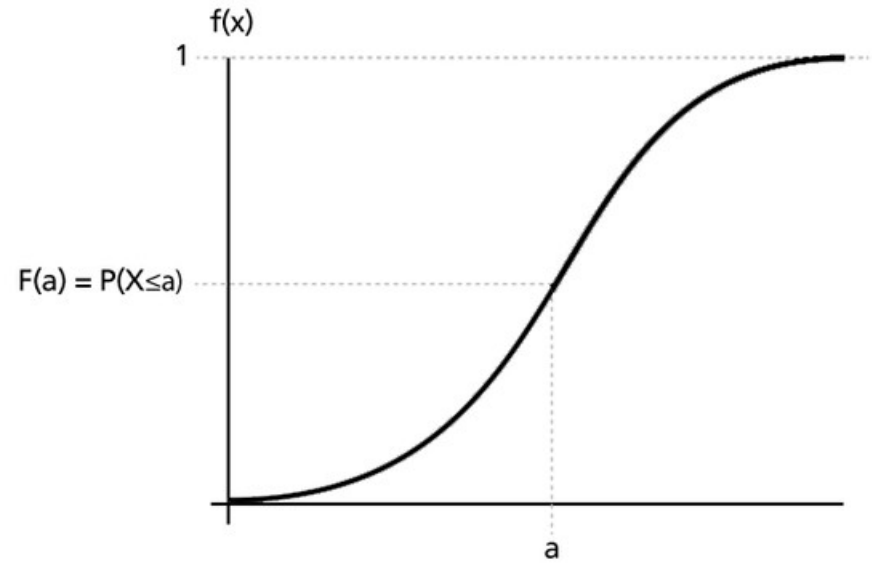
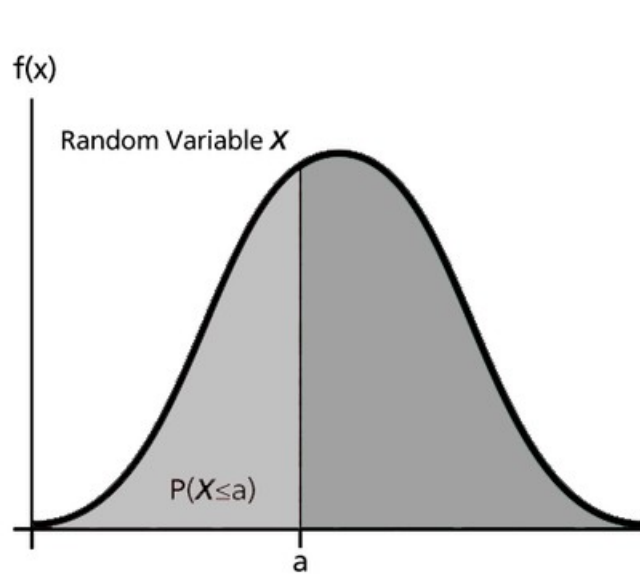


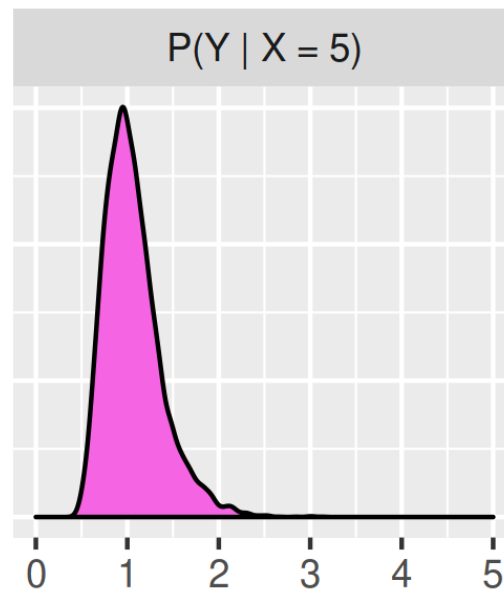
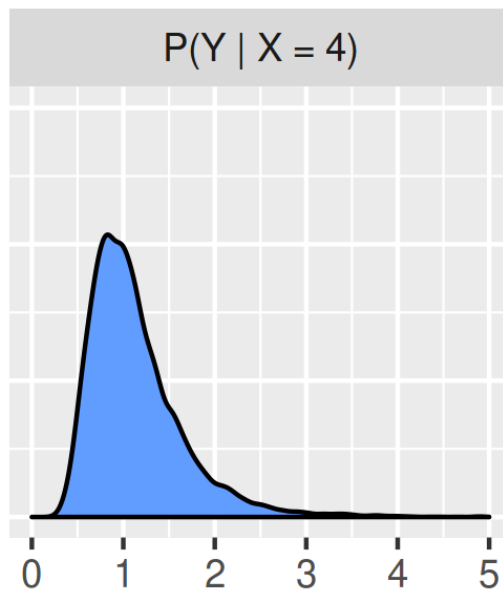
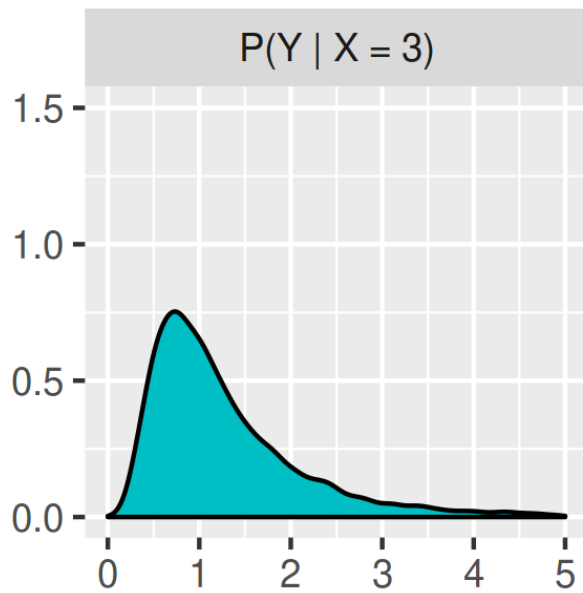
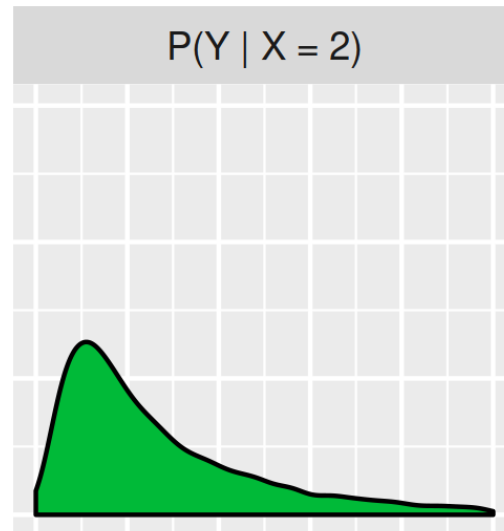
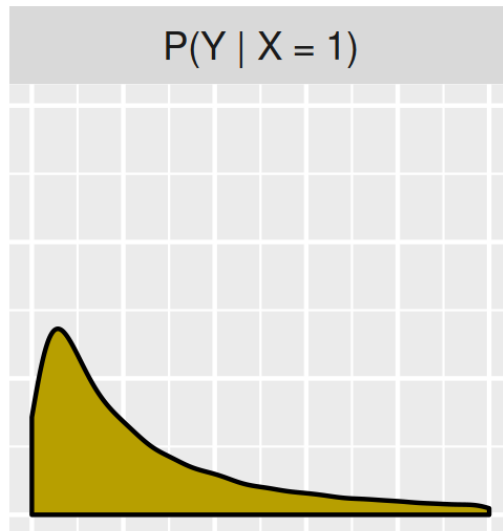
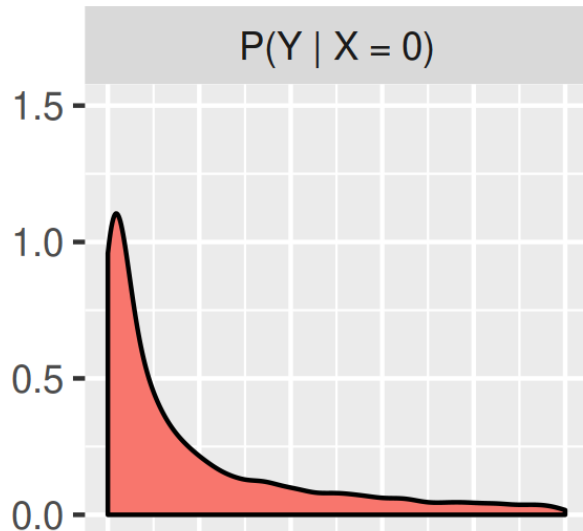
Probabilidades y densidades condicionales



Recordemos

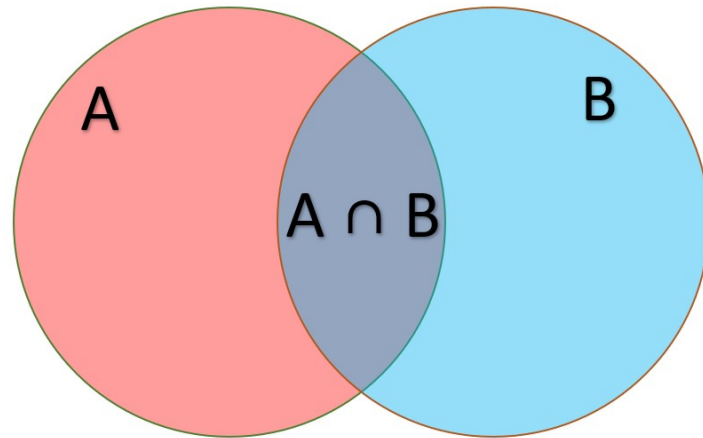
Densidad vs Acumulada





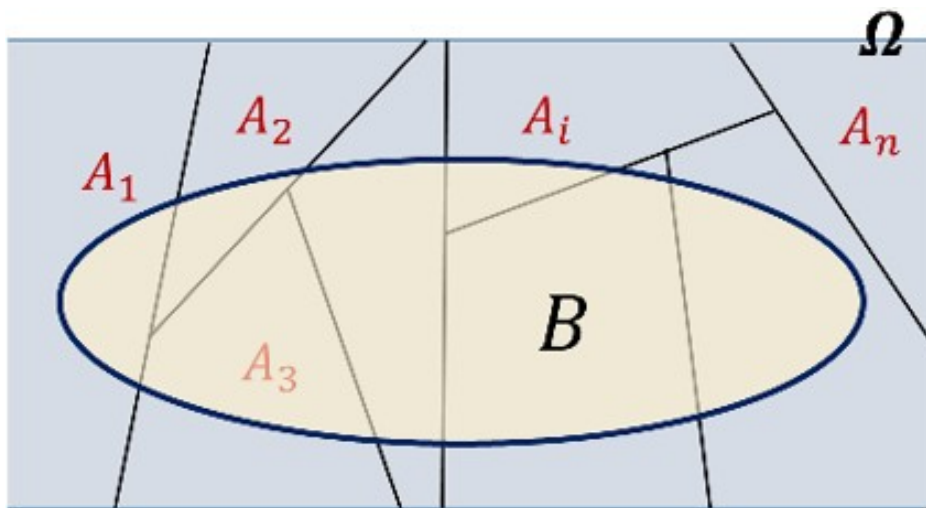
Regla multiplicativa

$$P(B \cap A) = P(B) * P(A|B) \text{ o bien } P(A|B) = \frac{P(B)}{P(B \cap A)}$$



Ley de probabilidad total

$$P(B) = \sum_{i=1}^n P(B \cap A_i) = \sum_{i=1}^n P(B|A_i)P(A_i)$$



Teorema de Bayes

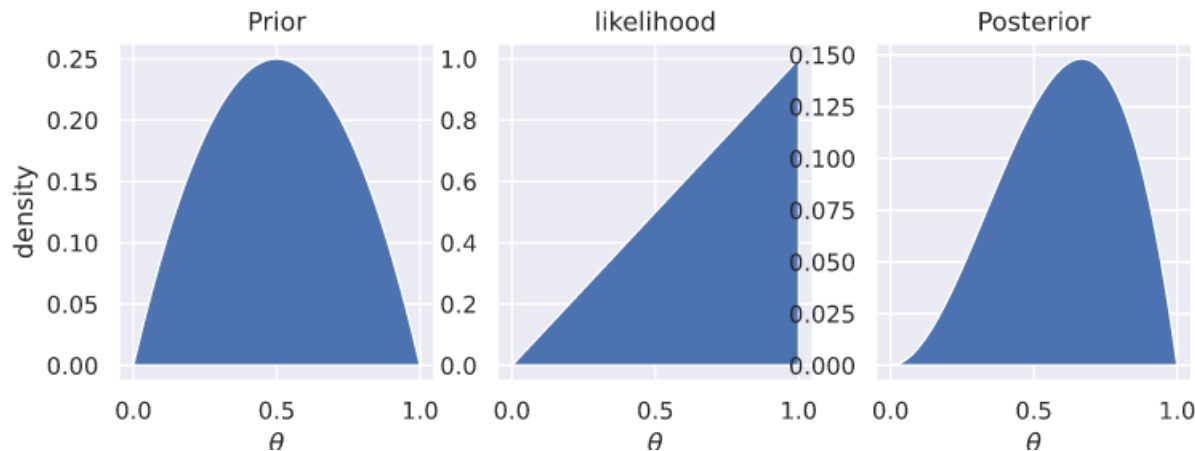
$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)}$$

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{k=1}^n P(B|A_k)P(A_k)}$$

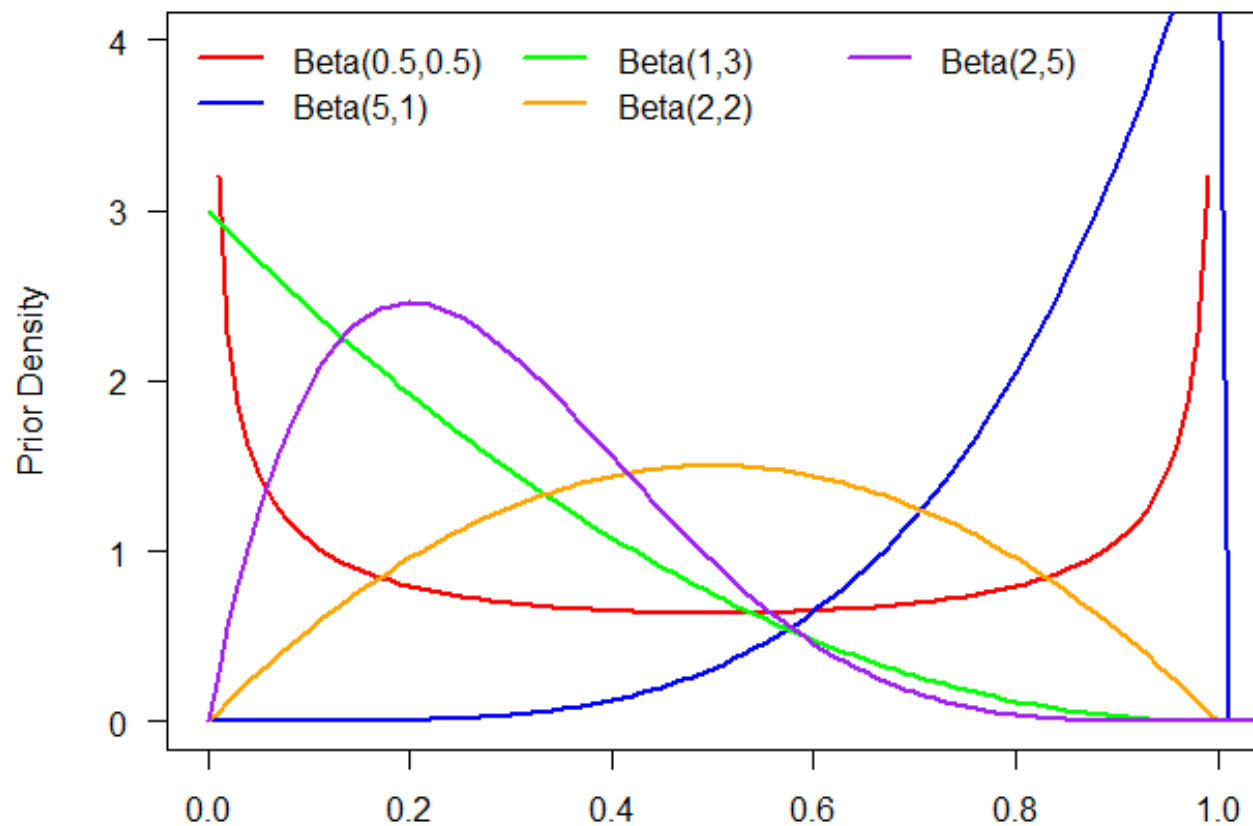
Teorema de Bayes para actualizar la previa

$$P(\theta|X) \propto P(\theta) * P(X|\theta)$$

$$f(\theta|X) \propto f(\theta) * f(X|\theta) = f(\theta) * L(\theta; X)$$



Prior Distributions



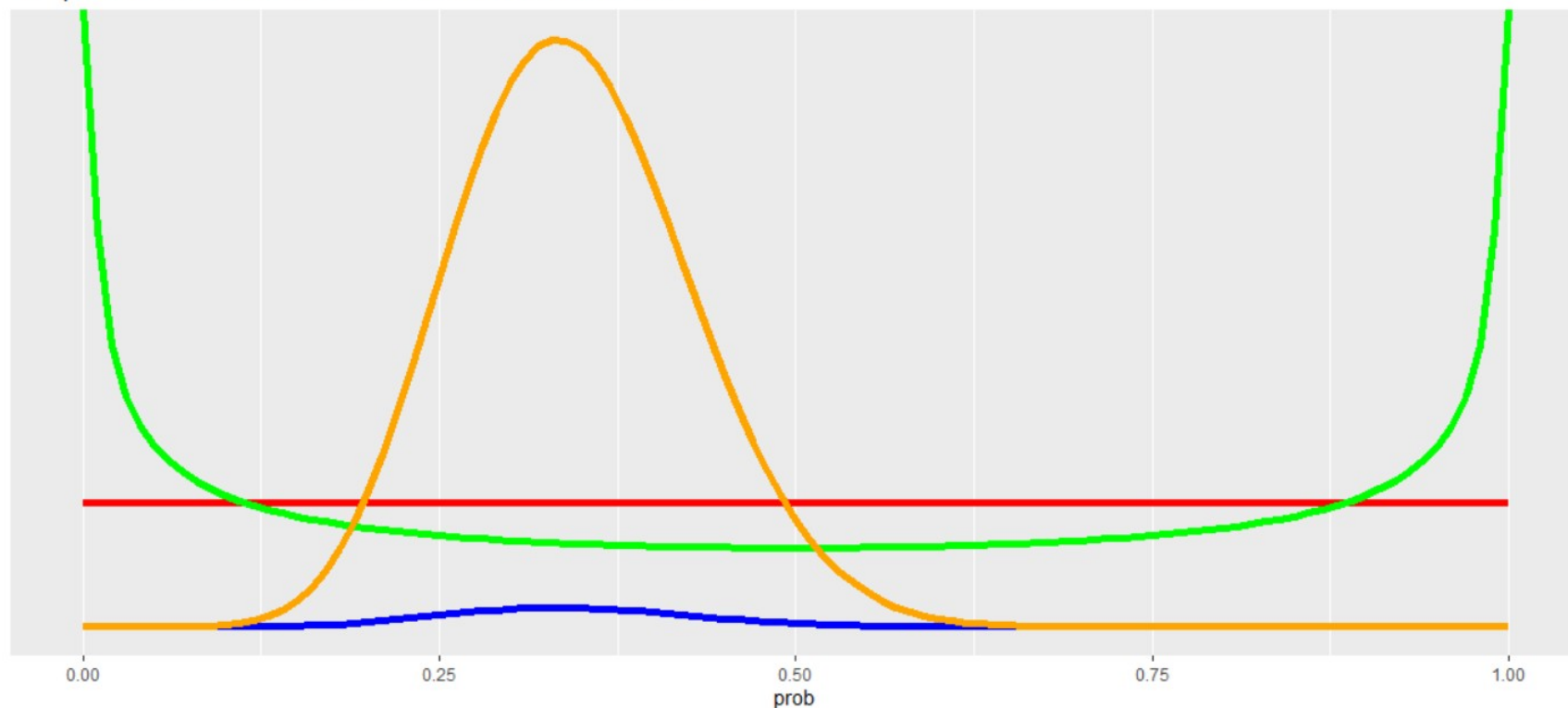
$$f(p) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1} (1-p)^{\beta-1}$$

$$\begin{aligned} p(\pi|y) &\propto p(y|\pi)p(\pi) \\ &= \text{Binomial}(n, \pi) \times \text{Beta}(\alpha, \beta) \\ &= \binom{n}{y} \pi^y (1-\pi)^{(n-y)} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \pi^{\alpha-1} (1-\pi)^{\beta-1} \\ &\propto \pi^y (1-\pi)^{(n-y)} \pi^{\alpha-1} (1-\pi)^{\beta-1} \end{aligned}$$

$$p(\pi|y) \propto \pi^{y+\alpha-1} (1-\pi)^{n-y+\beta-1}$$

Mean	$\alpha/(\alpha + \beta)$
Median	$\approx (\alpha - 1/3)/(\alpha + \beta - 1/3)$, for $\alpha, \beta > 1$
Mode	$(\alpha - 1)/(\alpha + \beta - 2)$, for $\alpha, \beta > 1$
Variance	$\alpha\beta/[(\alpha + \beta)^2(\alpha + \beta + 1)]$
Skewness	$2(\beta - \alpha)\sqrt{\alpha + \beta + 1}/[(\alpha + \beta + 2)\sqrt{\alpha\beta}]$
Kurtosis	$6[(\beta - \alpha)^2(\alpha + \beta + 1) - \alpha\beta(\alpha + \beta + 2)] / [\alpha\beta(\alpha + \beta + 2)(\alpha + \beta + 3)]$

Componentes de la verosimilitud caso Volados o Prevalencia



```
stat_function(fun = dbeta,size=2, n = 101,color="red", args = list(shape1=1, shape2=1)) +  
stat_function(fun = dbeta,size=2, n = 101,color="green", args = list(shape1=1/2, shape2=1/2))  
stat_function(fun = dbinom,size=2, n = 101,color="blue", args = list(x = 10, size=30))+  
stat_function(fun = dbeta,size=2, n = 101,color="orange", args = list(shape1=11, shape2=21))+
```