Seja $(X_1, \dots, X_n)$ uma amostra aleatória (a.a) de $X \sim f_{\theta}, \theta \in \Theta$

a Função Densidade de Probabilidade conjunta de $(X_1, X_2, \dots, X_n)$ é $\forall \theta \in \Theta$,
no caso discreto:

$$P_\theta(X_1 = k_1, X_2 = k_2, \dots, X_n = k_n) \overset{\text{ind}}{=} \prod_{i=1}^{n} P_\theta(X_i = k_i) \overset{\text{id}}{=} \prod_{i=1}^{n} P_\theta(X = k_i) \overset{\text{def}}{=} \prod_{i=1}^{n} f_\theta(k_i)$$

no caso contínuo:

$$f_\theta^{(x)}(k_1, k_2, \dots, k_n) \overset{\text{i.i.d}}{=} \prod_{i=1}^{n} f_\theta(k_i)$$

Exemplo

Seja (X_1, X_2, X_3) uma a.a de $X \sim \text{Ber}(\theta), \theta \in (0,1)$

1. Especifique o espaço paramétrico
2. Calcule a função de probabilidade da amostra
3. Encontre as seguintes Quantidade de Interesse em função de θ
    1. $g(\theta) = E_\theta(X)$
    2. $g(\theta) = P_\theta(X = 0)$
    3. $g(\theta) = \text{CV}_\theta(X)$
    Resolução
4. $\Theta = (0, 1)$
5. Duas resoluções possíveis
    1. Dando valores à amostra$$\begin{aligned}
       &(X{1}, X{2},X{3}) &P(X{1}= k{1},X{2}= k{2},X{3}= k{3})= \prod ^{3}{i=1}P\theta(X=k{1})\
       &(0,0,0) &(1-\theta)^{3}\
       &(0,0,1) &(1-\theta)^{2}\theta\
       &(0,1,0) &(1-\theta)^{2}\theta\
       &(1,0,0) &(1-\theta)^{2}\theta\
       &(0,1,1) &(1-\theta)\theta^2\
       &(1,0,1) &(1-\theta)\theta^{2}\
       &(1,1,0) &(1-\theta)\theta^{2}\
       &(1,1,1) &\theta^{3}\ \end{aligned}

2.EnunciandoafunçãoObserveque, se$k ∈ {0,1}$,

```
\begin{aligned}
&P\theta(X=k)=\theta^{k}(1-\theta)^{1-k}\cdot\mathbb{1}
{{0,1 }}(k)\Rightarrow\
&P\theta(X{1}=k,X{2}=k{2},X{3}=
k{3})\stackrel{\text{i.i.d}}{=}\prod^{3}{i=1}
{\theta^{k{i}}(1-\theta)^{1-k{1}}\mathbb{1}{{0,1 }}
(k{1}) } = \
&= \theta^{\sum\limits^{3}{i=1}k{1}}(1-\theta)^{3-
\sum\limits^{3}{i=1}k{1}}\prod^{3}{i=1}\mathbb{1}{{0,1}}
(k{i})
\end{aligned}$$
```

6. Em função de θ:
   1. $g(\theta) = E_\theta(X) = \theta$
   2. $g(\theta) = P_\theta(X=0) = 1-\theta$
   3. $g(\theta) = \mathrm{CV}_\theta(X) = \frac{\sqrt{\theta(1-\theta)}}{\theta}$

Outro exemplo
Seja $(X_1, \ldots, X_n)$ uma a.a de $X \sim \mathrm{Ber}(\theta), \theta \in (0,1)$, encontre a f.p conjunta da amostra.

$$P_\theta(X_1 = k_1, \ldots, X_n = k_n) \stackrel{\text{iid}}{=} \prod_{i=1}^{n} P_\theta(X = k_i) = \prod_{i=1}^{n} \{\theta^{k_i}(1-\theta)^{1-k_i}\mathbb{1}_{\{0,1\}}(k_i)\}$$

$$\Rightarrow P_\theta(X_1 = k_1, \ldots, X_n = k_n) = \theta^{\sum\limits_{i=1}^{n} k_1} \cdot \theta^{n - \sum\limits_{i=i}^{n} k_i} \cdot \mathbb{1}_{\{0,1\}}(k_i)$$

Mais um exemplo
Seja $(X_1, \ldots, X_n)$ uma a.a de $X \sim \mathrm{Pois}(\theta), \theta \in (0,\infty)$, encontre a f.p. conjunta da amostra.
Como esse vetor é uma a.a. (ou seja, variáveis aleatórias independentes e identicamente distribuídas), temos que

$$P_\theta(X_1 = k_1, \ldots, X_n = k_n) \stackrel{\text{iid}}{=} \prod_{i=1}^{n} P_\theta(X = k_i) = \prod_{i=1}^{n} \{\mathrm{e}^{-\theta} \cdot \frac{\theta^{k_i}}{k_i!}\}$$

Sempre que $k_i \in \{0, 1, \ldots\}, \forall i = 1, \ldots, n$

$$\Rightarrow P_\theta(\ldots) = \mathrm{e}^{-n\theta} \cdot \frac{\theta^{\sum\limits_{i=1}^{n} k_i}}{\prod_{i=1}^{n}(k_i)!}$$

Um exemplo para contínua,
Seja $(X_1, \ldots, X_n)$ uma a.a de $X \sim \mathrm{Exp}(\theta), \theta \in (0,\infty)$, encontre a f.d.p

conjunta da amostra.

$$F_\theta^{(*)}(k_1,\ldots,k_n) \stackrel{\text{iid}}{=} \prod_{i=1}^{n} f_\theta(k_i) = \prod_{i=1}^{n}\{\theta \mathrm{e}^{-\theta k_i} \cdot \mathbb{1}_{(0,\infty)}(k_i)\} \Rightarrow f_\theta^{(*)}(\ldots) = \theta^n \cdot \mathrm{e}^{-\theta \sum_{i=1}^{n} k_i} \cdot \prod_{i=1}^{n} \mathbb{1}_{(0,\infty)}(k)$$

Mais um,
Seja $(X_1,\ldots,X_n)$ uma a.a. (i.i.d) de $X \sim N(\mu,\sigma^2)$ em que
$\theta = (\mu,\sigma^2) \in \Theta = \mathbb{R} \times \mathbb{R}_+$. Considere $\underset{\sim}{x} = (x_1,\ldots,x_n)$ a amostra observada.

$$L_{\underset{\sim}{X}} \stackrel{\text{iid}}{=} \prod_{i=1}^{n} f_\theta(x_i) = \prod_{i=1}^{n}\left\{\frac{1}{\sqrt{2\pi\sigma^2}}\exp\{-\frac{1}{2\sigma^2}(x_1-\mu)^2\}\right\} = \frac{1}{(2\pi\sigma^2)^{\frac{x}{2}}} \cdot \exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\}$$

Função de Verossimilhança

Obs: quando analisamos a distribuição conjunta da amostra em função de θ nos valores da amostra observada, temos a **Função de Verossimilhança**

$$\mathrm{L}_{\underset{\sim}{x}}(\theta) = P_\theta(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$

em que $\underset{\sim}{x} = (x_1,x_2,\ldots,x_n)$ é a amostra observada.
Obs: A função de verossimilhança, no caso discreto, é a probabilidade de observar a amostra observada.

Exemplo
Considere (X_1,X_2,X_3,X_4) a.a de $X \sim \mathrm{Ber}(\theta), \theta \in \{0.1,0.5,0.9\}$. Note que o espaço paramétrico é $\Theta = \{0.1,0.5,0.9\}$. Considere, ainda, que a amostra observada foi $(0,1,1,1)$. Encontre a função de verossimilhança.

$$L_{\underset{\sim}{X}}(\theta) = P_\theta(X_1 = x_1, X_2 = x_2, X_3 = x_3, X_4 = x_4) = \theta^{\sum_{i=1}^{4} x_i}(1-\theta)^{4-\sum_{i=1}^{4} x_i} \cdot \cancelto{1}{\prod_{i=1}^{4} \mathbb{1}_{\{0,1\}}(x_i)} = \theta^3(1$$

Estimação via máxima verossimilhança (EMV)

O valor numérico $\hat{\theta}_n$ que maximiza a função de verossimilhança, ou seja, $L_{\underset{\sim}{x}}(\hat{\theta}_n) \geq L_{\underset{\sim}{x}}(\theta) \forall \theta \in \Theta$ é dito ser uma estimativa de máxima verossimilhança (MV) para θ. Observe que $\hat{\theta}_n$ depende da amostra observada e portanto: $\hat{\theta}_n = \hat{\theta}_(x_1,x_2,\ldots,x_n)$.

O estimador de máxima verossimilhança é obtido substituindo $(x_1,\ldots,x_n)$ por $(X_1,\ldots,X_n)$, ou seja, $\hat{\theta}_{(X_1,\ldots,X_n)}$ é o Estimador de

Máxima Verossimilhança (EMV)

Exemplo:

Seja $(X_1, \ldots, X_n)$ a.a de $X \sim f_\theta, \theta \in \{\frac{1}{3}, \frac{1}{2}\}$ em que f_θ é uma função de probabilidade que satisfaz:

$X = x:$	0	1	2
$f_\theta(x):$	θ	θ^2	$1-\theta-\theta^2$

Considere que a amostra observada é $\underset{\sim}{x} = (0, 0, 1)$.

a-) Encontre a estimativa da máxima verossimilhança
Sabemos que

$$f_\theta(x) = \theta^{\mathbb{1}_{\{0\}}(x)} \cdot (\theta^2)^{\mathbb{1}_{\{1\}}(x)} \cdot (1-\theta-\theta^2)^{\mathbb{1}_{\{2\}}(x)} \forall \theta \in \Theta$$

portanto,

$$L_{\underset{\sim}{x}}(\theta) \overset{\text{iid}}{=} \prod_{i=1}^{n} f_\theta(x_i) = \theta^{\sum\limits_{i=1}^{n} \mathbb{1}_{\{0\}}(x_i)} \cdot (\theta^2)^{\sum\limits_{i=1}^{n} \mathbb{1}_{\{1\}}(x_i)} \cdot (1-\theta-\theta^2)^{\sum\limits_{i=1}^{n} \mathbb{1}_{\{2\}}(x_i)} \forall \theta \in \Theta$$

Para $\underset{\sim}{x} = (0, 0, 1)$,

$$L_{\underset{\sim}{x}}(\theta) = \theta^2 \cdot \theta^2 \cdot (1-\theta-\theta^2)^0 = \theta^4 \forall \theta \in \Theta$$

Substituindo $\forall \theta \in \Theta$:

$$\theta = \frac{1}{2} \Rightarrow L_{\underset{\sim}{x}}\left(\frac{1}{2}\right) = \frac{1}{16} \quad \theta = \frac{1}{3} \Rightarrow L_{\underset{\sim}{x}}\left(\frac{1}{3}\right) = \frac{1}{81}$$

Portanto, $\hat{\theta}_n = \frac{1}{2}$ é a estimativa de máxima verossimilhança.

Invariância dos EMVs

Teorema.
Se $\hat{\theta}_{(X_1, \ldots, X_n)}$ for EMV para θ, então $g(\hat{\theta}_{(X_1, \ldots, X_n)})$ é o EMV para $g(\theta)$, ou seja, $g(\hat{\theta}_n)$ é a estimativa de máxima verossimilhança para $g(\theta)$

Mais um exemplo.

Seja $(X_1, \ldots, X_n)$ a.a. de $X \sim N(\mu, \sigma^2)$ em que $\theta = (\mu, \sigma^2) \in \Theta = \mathbb{R} \times \mathbb{R}^+$
Assuma que $\underset{\sim}{x} = (x_1, \ldots, x_n)$ é a amostra observada.
Lembrando que estaremos chamando $\theta = (\mu, \sigma^2)$, mas estes são parâmetros genéricos. Poderíamos, por exemplo, chamá-los de $\theta = (\theta_1, \theta_2)$, o que pode facilitar a visualizar algumas derivadas.

a-) Encontre as estimativas de máxima verossimilhança para $\theta = (\mu, \sigma^2)$:

A Função de Verossimilhança é:

$$
\begin{aligned}
L_{\underset{\sim}{x}}(\theta) &\overset{\text{iid}}{=} \prod_{i=1}^{n} f_\theta(x_i) = \prod_{i=1}^{n} \left\{ \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{ \frac{-1}{2} \cdot \frac{(x_i - \mu)^2}{\sigma^2} \right\} \right\} \\
&= \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \cdot \exp\left\{ \frac{-1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 \right\}
\end{aligned}
$$

Podemos derivar para encontrar o máximo da FMV. Para isso, derivaremos e igualamos a zero primeiro em relação a μ e então a σ^2 (podemos aplicar o logaritmo para facilitar as operações.)

$$
\begin{aligned}
\frac{\partial \ln(L_{\underset{\sim}{x}})}{\partial \mu} &= \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu) = 0 \Rightarrow \hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i \\
\frac{\partial \ln(L_{\underset{\sim}{x}})}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (x_i - \mu)^2 = 0 \\
&\therefore \\
\text{Estimativas MV} &= \begin{cases} \mu = \bar{x} \\ -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum\limits_{i=1}^{n} (x_i - \mu)^2 = 0 \end{cases} \\
&\Leftrightarrow \begin{cases} \hat{\mu} = \bar{x} \\ \hat{\sigma}^2 = \frac{1}{n} \sum\limits_{i=1}^{n} (x_i - \bar{x})^2 \end{cases}
\end{aligned}
$$

Estes são os pontos que maximizam a Função de Máxima Verossimilhança. (Provados em cálculo), ou seja, são as estimativas de máxima verossimilhança para μ, σ^2 respectivamente, e $\hat{\mu}(X_1, \dots, X_n) = \bar{X}, \sigma^2(X_1, \dots, X_n) = \frac{1}{n} \sum\limits_{i=1}^{n} (X_i - \bar{X})^2$ são os estimadores de máxima verossimilhança.

Pela propriedade de invariância podemos encontrar o EMV para $g(\theta) = \frac{\sqrt{\text{Var}_\theta(X)}}{E_{\theta(X)}}$:

$$
\widehat{g(\theta)} = \frac{\sqrt{\frac{1}{n} \sum\limits_{i=1}^{n} (X_i - \bar{X})}}{\bar{X}}
$$

Observação:

Seja $(X_1, \dots, X_n)$ a.a. de $X \sim N(\mu, \sigma^2)$. Então,

1-) $\bar{X} \underset{\text{Exata!}}{\sim} N\left(\mu, \frac{\sigma^2}{n}\right) \forall \mu, \sigma \in \mathbb{R} : \sigma^2 > 0 \text{ e } n \geq 1$

2-) $\sum_{i=1}^{n} \frac{(x_1-\bar{X})^2}{\sigma^2} \underset{\text{Exata!}}{\sim} \chi^2_{(n-1)}$ em que χ^2_k representa a Distribuição Qui-Quadrado com k grau de liberdade, cuja função densidade de probabilidade é:

$$f(x) = \frac{1}{\Gamma(\frac{k}{2})2^{\frac{k}{2}}} \cdot x^{\frac{k}{2}-1} \cdot \exp\left\{\frac{-x}{2}\right\} \cdot \mathbb{1}_{(0,\infty)}(x)$$

Para qualquer outra distribuição, existe um resultado aproximado pelo

Teorema do Limite Central

Seja $X_1, \dots, X_n$ a.a de $X \sim f_\theta, \theta \in \Theta : E_\theta(X^2) < \infty$, então:

$$\bar{X} \overset{\text{Aproximadamente}}{\sim} N\left(E_\theta(X), \frac{Var_\theta X}{n}\right)$$

Formalmente, temos o enunciado do *Teorema do Limite Central*:

$$\frac{\sqrt{n}(\bar{X} - E_\theta(x))}{\sqrt{\text{Var}_\theta(X)}} \overset{\text{Distribuição}}{\underset{n\to\infty}{\longrightarrow}} N \sim (0,1) \forall \theta \in \Theta$$

Se $X \sim N(\mu, \sigma)$, então a distribuição é exata.

Ademais, seja g uma função contínua e diferenciável tal que $g'(\theta) \neq 0$. Então,

$$g(\bar{X}) \overset{\text{Aproximadamente}}{\sim} N\left(g(E_\theta(X)), \frac{g'(E_\theta(X))^2 \text{Var}_\theta(X)}{n}\right)$$

Exemplo:

Seja $(X_1, \dots, X_n)$ a.a de $X \sim Ber(\theta), \theta \in (0,1)$. Já vimos que EMV p/θ é

$$

\bar{\theta}(X{1},\dots,X{n}) = \bar{X}

$$eoEMVp/\$g(\theta) = \text{Var}_{\theta(x)} = \theta(1-\theta)\$é:$$

\widehat{g(\theta)} = \bar{X}(1-\bar{X}).

$$\bar{X} \overset{\text{approx.}}{\sim} N\left(\theta, \frac{\theta(1-\theta)}{n}\right)$$

$$Agora, \$\$ g(\bar{X}) = \bar{X}(1-\bar{X}) \overset{\text{approx}}{\sim} N\left(\theta(1-\theta), \frac{[g'(\theta)]^2\theta(1-\theta)}{n}\right)$$

$$\Rightarrow N\left(\theta(1-\theta), \frac{(1-2\theta)^2\theta(1-\theta)}{n}\right)$$