FORMULÁRIO DE ESTATÍSTICA I

PROBABILIDADE

- $P(A \cap B \cap C) = P(A)P(B \mid A)P(C \mid A \cap B)$
- Sendo $\{A_1, A_2,..\}$ uma partição do espaço dos resultados com $P(A_j) > 0$, $j = 1,2,...$,

$$P(B) = \sum_j P(A_j)P(B \mid A_j) \quad ; \quad P(A_j \mid B) = \frac{P(A_j)P(B \mid A_j)}{\sum_i P(A_i)P(B \mid A_i)}.$$

VALOR ESPERADO, MOMENTOS E PARÂMETROS

	Discretas	Contínuas
$E[\psi(X)] =$	$\sum_x \psi(x) f_X(x)$	$\int_{-\infty}^{+\infty} \psi(x) f_X(x)\, dx$
$E[\psi(X,Y)] =$	$\sum_x \sum_y \psi(x,y) f_{X,Y}(x,y)$	$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \psi(x,y) f_{X,Y}(x,y)\, dx\, dy$
$E[\psi(X,Y) \mid X = x] =$	$\sum_y \psi(x,y) f_{Y\mid X=x}(y)$	$\int_{-\infty}^{+\infty} \psi(x,y) f_{Y\mid X=x}(y) dy$

Momentos de ordem k	$\mu'_k = E(X^k)$	$\mu_k = E\,[\,(X-\mu)^k\,]$
Momentos de ordem *r*+*s*	$\mu'_{rs} = E(X^r Y^s)$	$\mu_{rs} = E\{(X-\mu_X)^r (Y-\mu_Y)^s\}$

$\text{Var}(X) = E(X-\mu)^2 = E(X^2) - \mu^2$;

$\text{Cov}(X,Y) = E\{(X-\mu_X)(Y-\mu_Y)\} = E(XY) - E(X)E(Y)$; $\rho_{X,Y} = \dfrac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}$

$\text{E}(aX+bY) = aE(X) + bE(Y)$ e $\text{Var}(aX+bY) = a^2\text{Var}(X) + b^2\text{Var}(Y) + 2ab\,\text{Cov}(X,Y)$ com a, b constantes

$E(Y) = E_X[E(Y \mid X)]$; $\text{Var}(Y) = \text{Var}_X[E(Y \mid X)] + E_X[\text{Var}(Y \mid X)]$

Coeficiente de assimetria: $\gamma_1 = \dfrac{\mu_3}{\sigma^3}$; *Kurtosis*: $\gamma_2 = \dfrac{\mu_4}{\sigma^4}$

Quantil (caso contínuo): $\xi_\alpha : \int_{-\infty}^{\xi_\alpha} f(x)dx = \alpha \Leftrightarrow F(\xi_\alpha) = \alpha$

Função geradora de momentos: $M_X(s) = E(e^{sX})$; $E(X^r) = M_X^{(r)}(0)$

DISTRIBUIÇÕES TEÓRICAS

- **UNIFORME (DISCRETA)**

Caso $f(x) = \dfrac{1}{n}$, $x = 1,2,...,n$; $E(X) = \dfrac{n+1}{2}$; $\text{Var}(X) = \dfrac{n^2-1}{12}$

Caso $f(x) = \dfrac{1}{m+1}$, $x = 0,1,2,...,m$; $E(X) = \dfrac{m}{2}$; $\text{Var}(X) = \dfrac{m(m+2)}{12}$

- **BINOMIAL** $X \sim B(n;\theta)$, $(0 < \theta < 1)$

$f(x \mid \theta) = \dbinom{n}{x} \theta^x (1-\theta)^{n-x}$, $x = 0, 1, 2, ..., n$

$E(X) = n\theta$; $\text{Var}(X) = n\theta(1-\theta)$; $M_X(s) = \left[(1-\theta) + \theta e^s\right]^n$; $\gamma_1 = (1-2\theta)/\sigma$

Propriedades:

- $X \sim B(n;\theta) \Leftrightarrow (n-X) \sim B(n;1-\theta)$
- $X_1 \sim B(n_1;\theta,)$, $X_2 \sim B(n_2;\theta)$, X_1 e X_2 independentes $\Rightarrow X_1 + X_2 \sim B(n_1+n_2,\theta)$

- **BERNOULLI** $X \sim B(1;\theta)$

- **POISSON** $X \sim \text{Po}(\lambda)$, $(\lambda > 0)$

  $f(x \mid \lambda) = \dfrac{e^{-\lambda}\lambda^x}{x!}$, $x = 0,1,2,...$; $E(X) = \lambda$; $\text{Var}(X) = \lambda$; $M_X(s) = \exp\{\lambda(e^s - 1)\}$; $\gamma_1 = \lambda^{-1/2}$

  Propriedades:

  - $X_1 \sim \text{Po}(\lambda_1)$, $X_2 \sim \text{Po}(\lambda_2)$, X_1 e X_2 independentes $\Rightarrow X_1 + X_2 \sim \text{Po}(\lambda_1 + \lambda_2)$
  - Se $X \sim B(n;\theta)$, com n grande θ pequeno então $X \overset{a}{\sim} \text{Po}(n\theta)$

- **UNIFORME (CONTÍNUA)** $X \sim U(\alpha, \beta)$, $(\alpha < \beta)$

  $f(x \mid \alpha, \beta) = \dfrac{1}{\beta - \alpha}$ $\alpha < x < \beta$; $E(X) = \dfrac{\alpha + \beta}{2}$; $\text{Var}(X) = \dfrac{(\beta - \alpha)^2}{12}$; $M_X(s) = \dfrac{e^{s\beta} - e^{s\alpha}}{s(\beta - \alpha)}$, $s \neq 0$

- **NORMAL** $X \sim N(\mu, \sigma^2)$, $(-\infty < \mu < +\infty, \quad 0 < \sigma < +\infty)$

  $f(x \mid \mu, \sigma^2) = \dfrac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\dfrac{1}{2\sigma^2}(x - \mu)^2\right\}$, $-\infty < x < +\infty$

  $E(X) = \mu$; $\text{Var}(X) = \sigma^2$; $M_X(s) = \exp\left\{\mu s + \dfrac{\sigma^2 s^2}{2}\right\}$; $\gamma_1 = 0$; $\gamma_2 = 3$

  Propriedades:

  - Normal estandardizada $Z = \dfrac{X - \mu}{\sigma} \sim N(0,1)$; $\phi(z) = \phi(-z)$ e $\Phi(z) = 1 - \Phi(-z)$
  - $X_i \sim N(\mu, \sigma^2)$ ($i = 1, 2, ..., k$) independentes $\Rightarrow Y = \sum_{i=1}^{k} X_i \sim N(k\mu, k\sigma^2)$ e $\overline{X} = \dfrac{1}{k}\sum_{i=1}^{k} X_i \sim N\left(\mu, \dfrac{\sigma^2}{k}\right)$
  - $X_i \sim N(\mu_i, \sigma_i^2)$ ($i = 1, 2, ..., k$) independentes $\Rightarrow \sum_{i=1}^{k} \alpha_i X_i \sim N(\mu_Y, \sigma_Y^2)$ com $\mu_Y = \sum_{i=1}^{k} \alpha_i \mu_i$ e $\sigma_Y^2 = \sum_{i=1}^{k} \alpha_i^2 \sigma_i^2$

- **EXPONENCIAL** $X \sim \text{Ex}(\lambda)$, $(\lambda > 0)$; $X \sim \text{Ex}(\lambda) \Leftrightarrow X \sim \text{G}(1, \lambda)$

  $f(x \mid \lambda) = \lambda e^{-\lambda x}$, $x > 0$ $E(X) = \dfrac{1}{\lambda}$; $\text{Var}(X) = \dfrac{1}{\lambda^2}$; $M_X(s) = \dfrac{\lambda}{\lambda - s}$, $s < \lambda$; $\gamma_1 = 2$; $\gamma_2 = 9$

  Propriedades:

  - $X_i \sim \text{Ex}(\lambda)$ ($i = 1, 2, ..., k$) independentes $\Rightarrow \sum_{i=1}^{k} X_i \sim G(k, \lambda)$ e $\min_i X_i \sim \text{Ex}(k\lambda)$

- **GAMA** $X \sim G(\alpha, \lambda)$, $(\lambda > 0, \alpha > 0)$

  $f(x \mid \alpha, \lambda) = \dfrac{\lambda^\alpha e^{-\lambda x} x^{\alpha - 1}}{\Gamma(\alpha)}$, $x > 0$; $E(X) = \dfrac{\alpha}{\lambda}$; $\text{Var}(X) = \dfrac{\alpha}{\lambda^2}$; $M_X(s) = \left(\dfrac{\lambda}{\lambda - s}\right)^\alpha$, $s < \lambda$; $\gamma_1 = \dfrac{2}{\sqrt{\alpha}}$; $\gamma_2 = 3 + \dfrac{6}{\alpha}$

  Propriedades:

  - $X_i \sim G(\alpha_i; \lambda)$, ($i = 1, 2, ..., k$) independentes $\Rightarrow \sum_{i=1}^{k} X_i \sim \text{G}\left(\sum_{i=1}^{k} \alpha_i ; \lambda\right)$
  - $X \sim G(\alpha, \lambda)$ então $Y = cX \sim G(\alpha, \dfrac{\lambda}{c})$ onde c constante positiva

- **QUI-QUADRADO** $X \sim \chi^2(n)$, (n inteiro positivo).

  $X \sim \chi^2(n) \Leftrightarrow X \sim \text{G}(n/2; 1/2)$; $E(X) = n$; $\text{Var}(X) = 2n$; $M_X(s) = (1 - 2s)^{-\frac{n}{2}}$, $s < \dfrac{1}{2}$; $\gamma_1 = \sqrt{\dfrac{8}{n}}$; $\gamma_2 = 3 + \dfrac{12}{n}$

  Propriedades:

  - $X_i \sim \chi^2_{(n_i)}$ ($i = 1, 2, ..., k$) independentes $\Rightarrow \sum_{i=1}^{k} X_i \sim \chi^2_{(n)}$ com $n = \sum_{i=1}^{k} n_i$
  - $X \sim \text{G}(n; \lambda) \Leftrightarrow 2\lambda X \sim \chi^2(2n)$
  - $X_i \sim N(0,1)$, ($i = 1,2,...,n$) independentes $\Rightarrow \sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$
  - $X \sim \chi^2(n) \Rightarrow \sqrt{2X} - \sqrt{2n - 1} \overset{a}{\sim} N(0,1)$

- ***t*-"STUDENT"**

$$T = \frac{U}{\sqrt{V/n}} \sim t(n) \text{ com } U \sim N(0,1) \text{ e } V \sim \chi^2(n) \text{ independentes}$$

$$E(T) = 0\,;\; \mathrm{Var}(T) = \frac{n}{n-2}\; (n>2)\,;\;\; \gamma_1 = 0\,;\; \gamma_2 = \frac{3(n-2)}{n-4}\; (n>4)$$

Propriedade: • Sendo $T \sim t(n) \Rightarrow \lim_{n\to\infty} F_T(t \mid n) = \Phi(t)$

- ***F*-SNEDCOR**

$$F = \frac{U/m}{V/n} \sim F(m,n) \text{ com } U \sim \chi^2(m),\; V \sim \chi^2(n) \text{ (independentes)}$$

$$E(X) = \frac{n}{n-2}\; (n>2)\,;\; \mathrm{Var}(X) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}\; (n>4)$$

Propriedades: • $X \sim F(m,n) \Rightarrow \frac{1}{X} \sim F(n,m)$ • $T \sim t_{(n)} \Rightarrow T^2 \sim F(1,n)$

TEOREMA DO LIMITE CENTRAL E COROLÁRIOS

TLC: Sendo X_i iid com $E(X_i) = \mu$ e $\mathrm{Var}(X_i) = \sigma^2 \;\Rightarrow\; \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\,\sigma} = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \overset{a}{\sim} N(0,1)$

Corolário: Sendo $X_i \sim B(1;\theta)$, iid $\Rightarrow \frac{\sum_{i=1}^{n} X_i - n\theta}{\sqrt{n\theta(1-\theta)}} \overset{a}{\sim} N(0,1)$

Correcção de continuidade: $P(a \le X \le b) \approx \Phi\left(\frac{b+\frac{1}{2}-n\theta}{\sqrt{n\theta(1-\theta)}}\right) - \Phi\left(\frac{a-\frac{1}{2}-n\theta}{\sqrt{n\theta(1-\theta)}}\right)$, com a e b inteiros

Corolário: Sendo $X \sim \mathrm{Po}(\lambda)$, quando $\lambda \to +\infty \;\Rightarrow\; \frac{X-\lambda}{\sqrt{\lambda}} \overset{a}{\sim} N(0,1)$

Correcção de continuidade: $P(a \le X \le b) \approx \Phi\left(\frac{b+\frac{1}{2}-\lambda}{\sqrt{\lambda}}\right) - \Phi\left(\frac{a-\frac{1}{2}-\lambda}{\sqrt{\lambda}}\right)$, com a e b inteiros

AMOSTRAGEM. DISTRIBUIÇÕES POR AMOSTRAGEM

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}\;;\quad S^2 = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n} = \frac{\sum_{i=1}^{n} X_i^2}{n} - \bar{X}^2\;;\quad (n-1)S'^2 = n\,S^2$$

$$E(\bar{X}) = \mu\;;\quad \mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n}\;;\quad E(S^2) = \frac{n-1}{n}\sigma^2\;;\quad E(S'^2) = \sigma^2$$

- **DISTRIBUIÇÃO DO MÍNIMO E DO MÁXIMO**

$$G_1(x) = 1 - [1-F(x)]^n \;;\quad G_n(x) = [F(x)]^n$$

- **GRANDES AMOSTRAS**

Caso geral

Média	$\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \overset{a}{\sim} N(0,1)$	$\dfrac{\overline{X}-\mu}{S'/\sqrt{n}} \overset{a}{\sim} N(0,1)$
Diferença de médias	$\dfrac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}} \overset{a}{\sim} N(0,1)$	$\dfrac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{S_1'^2}{m}+\dfrac{S_2'^2}{n}}} \overset{a}{\sim} N(0,1)$

População de Bernoulli

Proporção	$\dfrac{\overline{X}-\theta}{\sqrt{\dfrac{\theta(1-\theta)}{n}}} \overset{a}{\sim} N(0,1)$
Diferença de proporções	$\dfrac{\overline{X}_1-\overline{X}_2-(\theta_1-\theta_2)}{\sqrt{\dfrac{\theta_1(1-\theta_1)}{m}+\dfrac{\theta_2(1-\theta_2)}{n}}} \overset{a}{\sim} N(0,1)$

- **POPULAÇÕES NORMAIS**

Média	$\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$\dfrac{\overline{X}-\mu}{S'/\sqrt{n}} \sim t(n-1)$
Diferença de médias	$\dfrac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}} \sim N(0,1)$	$Z=\dfrac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{S_1'^2}{m}+\dfrac{S_2'^2}{n}}} \overset{a}{\sim} t(\nu)$ onde ν é o maior inteiro contido em r, $r=\dfrac{\left(\dfrac{s_1'^2}{m}+\dfrac{s_2'^2}{n}\right)^2}{\dfrac{1}{m-1}\left(\dfrac{s_1'^2}{m}\right)^2+\dfrac{1}{n-1}\left(\dfrac{s_2'^2}{n}\right)^2}$
	$T=\dfrac{\dfrac{\overline{X}_1-\overline{X}_2-(\mu_1-\mu_2)}{\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}}}{\sqrt{\dfrac{(m-1)S_1'^2+(n-1)S_2'^2}{m+n-2}}} \sim t(m+n-2)$	
Variância	$\dfrac{nS^2}{\sigma^2}=\dfrac{(n-1)S'^2}{\sigma^2} \sim \chi^2(n-1)$	
Relação de variâncias	$\dfrac{S_1'^2}{S_2'^2}\dfrac{\sigma_2^2}{\sigma_1^2} \sim F(m-1,n-1)$	