FORMULÁRIO DE ESTATÍSTICA II

• VALOR ESPERADO, MOMENTOS E PARÂMETROS

- $\text{Var}(X) = E(X-\mu)^2 = E(X^2) - \mu^2$
- $\text{Cov}(X,Y) = E\{(X-\mu_X)(Y-\mu_Y)\} = E(XY) - E(X)E(Y)$; $\rho_{X,Y} = \dfrac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}$
- $E(aX+bY) = aE(X) + bE(Y)$; $\text{Var}(aX+bY) = a^2\text{Var}(X) + b^2\text{Var}(Y) + 2ab\text{Cov}(X,Y)$ com a, b constantes

• DISTRIBUIÇÕES TEÓRICAS

• UNIFORME (DISCRETA)

- Caso $x = 1,2,\dots,n$: $f(x) = \dfrac{1}{n}$; $E(X) = \dfrac{n+1}{2}$; $\text{Var}(X) = \dfrac{n^2-1}{12}$
- Caso $x = 0,1,2,\dots,m$: $f(x) = \dfrac{1}{m+1}$; $E(X) = \dfrac{m}{2}$; $\text{Var}(X) = \dfrac{m(m+2)}{12}$

• BERNOULLI $X \sim B(1,\theta)$

$f(x\mid\theta) = \theta^x(1-\theta)^{1-x}$, $x = 0,1$ $(0<\theta<1)$; $E(X) = \theta$; $\text{Var}(X) = \theta(1-\theta)$

• BINOMIAL $X \sim B(n,\theta)$

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

$E(X) = n\theta$; $\text{Var}(X) = n\theta(1-\theta)$; $\Im(\theta) = \dfrac{n}{\theta(1-\theta)}$, n conhecido

Propriedades:

- $X \sim B(n,\theta) \Leftrightarrow (n-X) \sim B(n,1-\theta)$
- $X_i \sim B(n_i,\theta)$, independentes $(i = 1,2,\dots,k)$ $\Rightarrow \sum_{i=1}^{k} X_i \sim B(n,\theta)$, $n = \sum_{i=1}^{k} n_i$

• POISSON $X \sim \text{Po}(\lambda)$

$f(x\mid\lambda) = \dfrac{e^{-\lambda}\lambda^x}{x!}$, $x = 0,1,2,\dots$, $(\lambda>0)$; $E(X) = \lambda$; $\text{Var}(X) = \lambda$; $\Im(\lambda) = \dfrac{1}{\lambda}$

Propriedades:

- $X_i \sim \text{Po}(\lambda_i)$ independentes $(i = 1,2,\dots,k)$ $\Rightarrow X = \sum_{i=1}^{k} X_i \sim \text{Po}\left(\sum_{i=1}^{k}\lambda_i\right)$
- $X \sim B(n,\theta)$ com n grande e θ pequeno, então $X \overset{a}{\sim} \text{Po}(n\theta)$

• UNIFORME (CONTÍNUA) $X \sim U(\alpha,\beta)$

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

• NORMAL $X \sim N(\mu,\sigma^2)$

$f(x\mid\mu,\sigma^2) = \dfrac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\dfrac{1}{2\sigma^2}(x-\mu)^2\right\}$, $-\infty < x < +\infty$, $-\infty < \mu < +\infty$, $0 < \sigma < +\infty$;

$E(X) = \mu$; $\text{Var}(X) = \sigma^2$; $\Im(\mu) = \dfrac{1}{\sigma^2}$ σ^2 conhecido ; $\Im(\sigma^2) = \dfrac{1}{2\sigma^4}$ μ conhecido

Propriedades:

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

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

  $f(x\,|\,\lambda) = \lambda e^{-\lambda x}$ $x>0,\ \lambda>0$; $F(x\,|\,\lambda) = 1-e^{-\lambda x}$; $E(X)=\dfrac{1}{\lambda}$; $\text{Var}(X)=\dfrac{1}{\lambda^2}$; $\Im(\lambda)=\dfrac{1}{\lambda^2}$

  Propriedades:

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

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

  Função gama: $\Gamma(\alpha)=\int_0^{+\infty} e^{-x}x^{\alpha-1}dx$ $(\alpha>0)$ com $\Gamma(\alpha)=(\alpha-1)\Gamma(\alpha-1), \alpha>1$; $\Gamma(n)=(n-1)!$; $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$

  $f(x\,|\,\alpha,\lambda)=\dfrac{\lambda^{\alpha}e^{-\lambda x}x^{\alpha-1}}{\Gamma(\alpha)}$, $x>0,\ \alpha,\lambda>0$; $E(X)=\dfrac{\alpha}{\lambda}$; $\text{Var}(X)=\dfrac{\alpha}{\lambda^2}$; $\Im(\lambda)=\dfrac{\alpha}{\lambda^2}$ α conhecido

  Propriedades:

  - $X_i \sim G(\alpha_i,\lambda)$, $(i=1,2,...,k)$ independentes $\Rightarrow \sum_{i=1}^{k} X_i \sim \text{G}(\alpha,\lambda)$; $\alpha=\sum_{i=1}^{k}\alpha_i$
  - $X \sim G(\alpha,\lambda) \Rightarrow cX \sim G\left(\alpha,\dfrac{\lambda}{c}\right)$ $c>0$ constante

- **QUI-QUADRADO** $X \sim \chi^2(n)$

  $f(x\,|\,n)=\dfrac{e^{-\frac{x}{2}}x^{\frac{n}{2}-1}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}$, $x>0$, $n>0$ inteiro; $E(X)=n$; $\text{Var}(X)=2n$;

  Propriedades:

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

- ***t*-"STUDENT"** $T \sim t(n)$

  $T=\dfrac{U}{\sqrt{V/n}} \sim t(n)$ com $U \sim N(0,1)$ e $V \sim \chi^2(n)$ independentes; $E(T)=0$; $\text{Var}(T)=\dfrac{n}{n-2}$ $(n>2)$

- ***F*-SNEDCOR** $X \sim F(m,n)$

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

  $E(X)=\dfrac{n}{n-2}$ $(n>2)$; $\text{Var}(X)=\dfrac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$ $(n>4)$

  Propriedades: - $X \sim F(m,n) \Rightarrow \dfrac{1}{X} \sim F(n,m)$ - $T \sim t(n) \Rightarrow T^2 \sim F(1,n)$

- **TEOREMA DO LIMITE CENTRAL E COROLÁRIOS**

  X_i iid, com $E(X_i)=\mu$ e $\text{Var}(X_i)=\sigma^2 \Rightarrow \dfrac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\,\sigma}=\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \overset{a}{\sim} N(0,1)$

  Corolário: $X_i \sim B(1;\theta)$, independentes então $\dfrac{\sum_{i=1}^{n} X_i - n\theta}{\sqrt{n\theta(1-\theta)}} \overset{a}{\sim} N(0,1)$

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

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

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

- **AMOSTRAGEM. DISTRIBUIÇÕES POR AMOSTRAGEM**
  - **MÉDIA E VARIÂNCIA AMOSTRAIS**

$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$; $S^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2$; $S'^2 = \frac{n}{n-1}S^2$; $E(\overline{X}) = \mu$; $\mathrm{Var}(\overline{X}) = \frac{\sigma^2}{n}$; $E(S^2) = \frac{n-1}{n}\sigma^2$; $E(S'^2) = \sigma^2$

  - **DISTRIBUIÇÕES POR AMOSTRAGEM**

POPULAÇÕES NORMAIS

Média	$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$\frac{\overline{X}-\mu}{S'/\sqrt{n}} \sim t(n-1)$
Diferença de médias	Variâncias conhecidas $\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}} \sim N(0,1)$	
	Variâncias desconhecidas mas iguais $T = \frac{\overline{X}_1-\overline{X}_2-(\mu_1-\mu_2)}{\sqrt{\frac{1}{m}+\frac{1}{n}}\sqrt{\frac{(m-1)S_1'^2+(n-1)S_2'^2}{m+n-2}}} \sim t(m+n-2)$	
Variância	$\frac{nS^2}{\sigma^2} = \frac{(n-1)S'^2}{\sigma^2} \sim \chi^2(n-1)$	
Relação de variâncias	$\frac{S_1'^2}{S_2'^2}\frac{\sigma_2^2}{\sigma_1^2} \sim F(m-1,n-1)$ ou $\frac{S_2'^2}{S_1'^2}\frac{\sigma_1^2}{\sigma_2^2} \sim F(n-1,m-1)$	
Amostras emparelhadas	$\frac{\overline{Z}-\mu_X-\mu_Y}{S_Z'/\sqrt{n}} \sim t(n-1)$, (X_i,Y_i) - amostra emparelhada, $Z_i = X_i - Y_i$; $\overline{Z} = \overline{X} - \overline{Y}$	

GRANDES AMOSTRAS: CASO GERAL

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

GRANDES AMOSTRAS: POPULAÇÃO DE BERNOULLI

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

GRANDES AMOSTRAS: POPULAÇÃO DE POISSON

Média	$\dfrac{\overline{X}-\lambda}{\sqrt{\dfrac{\lambda}{n}}} \overset{a}{\sim} N(0,1)$	$\dfrac{\overline{X}-\lambda}{\sqrt{\dfrac{\overline{X}}{n}}} \overset{a}{\sim} N(0,1)$
Diferença de médias	$\dfrac{\overline{X}_1-\overline{X}_2-(\lambda_1-\lambda_2)}{\sqrt{\dfrac{\lambda_1}{m}+\dfrac{\lambda_2}{n}}} \overset{a}{\sim} N(0,1)$	$\dfrac{\overline{X}_1-\overline{X}_2-(\lambda_1-\lambda_2)}{\sqrt{\dfrac{\overline{X}_1}{m}+\dfrac{\overline{X}_2}{n}}} \overset{a}{\sim} N(0,1)$
Igualdade de médias	$\dfrac{\overline{X}_1-\overline{X}_2}{\sqrt{\left(\dfrac{1}{m}+\dfrac{1}{n}\right)\overline{X}}} \overset{a}{\sim} N(0,1)$ com $\overline{X}=\dfrac{m\overline{X}_1+n\overline{X}_2}{m+n}$	

- **ESTATÍSTICA-TESTE DO χ^2**

- **TESTE DE AJUSTAMENTO**

$Q=\sum_{j=1}^{m}\dfrac{(N_j-fe_j)^2}{fe_j} \overset{a}{\sim} \chi^2(m-1)$; $fe_j=np_{\circ j}$ - frequência esperada da classe j ;

Com estimação de k parâmetros para obter as estimativas $\hat{p}_{\circ j}$: $Q \overset{a}{\sim} \chi^2(m-k-1)$

- **TESTE DE INDEPENDÊNCIA**:

$Q=\sum_{i=1}^{r}\sum_{j=1}^{s}\dfrac{(N_{ij}-fe_{ij})^2}{fe_{ij}} \overset{a}{\sim} \chi^2((r-1)(s-1))$; $fe_{ij}=\dfrac{N_{i\circ}N_{\circ j}}{n}$ - frequência esperada da classe ij