1.3- Distribuições Não Centrais

1.3.1 - Distribuição Qui-Quadrado

A variável aleatória U tem distribuição Qui-Quadrado com n (inteiro, $n \geq 1$) graus de liberdade ($U \sim \chi^2_n$) se

$$f(u) = \frac{1}{\Gamma\left(\frac{n}{2}\right) 2^{n/2}} u^{\frac{n-2}{2}} e^{-\frac{u}{2}}, \qquad 0 < u < \infty$$

Caso particular da gama parâmetros $\frac{n}{2}$ e $\frac{1}{2}$

Obs:

$$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx \qquad a > 0$$

$\Gamma(a+1) = a\Gamma(a)$, a inteiro, $a = n$, $\Gamma(n+1) = n!$

Resultados

1- $m_U(t) = E(e^{tU}) = (1-2t)^{-n/2}$ $\qquad t < \frac{1}{2}$

função geradora de momentos de U

2- Momentos

$E(U) = n$ $\qquad Var(U) = 2n$

$E(U^2) = n(n+2)$ $\qquad E\left(\frac{1}{U}\right) = \frac{1}{n-2}$ $\quad n > 2$

$E\left(U^{1/2}\right) = \frac{2^{1/2}\,\Gamma\left(\frac{n+1}{2}\right)}{\Gamma(n/2)}$ $\qquad E\left(\frac{1}{U^2}\right) = \frac{1}{(n-2)(n-4)}$ $\quad n > 4$

$$E\left(U^{-1/2}\right) = \frac{\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{2}\,\Gamma\left(\frac{n}{2}\right)}$$

Teorema 1

Z_i, $i=1,2\dots n$ v.a. indep., $Z_i \sim N(0,1)$,

$$\sum_{i=1}^{n} Z_i^2 \sim \chi^2_n$$ soma de normais (0,1) ao quadrado

Teorema 2

Sejam $Y_1, Y_2 \dots Y_n$ amostra aleatória da dist. $N(\mu, \sigma^2)$.

Então

i) $$\sum_{i=1}^{n} \frac{(Y_i-\mu)^2}{\sigma^2} \sim \chi^2_n$$

ii) $$\sum_{i=1}^{n} \frac{(Y_i-\bar{Y})^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

iii) $\bar{Y}$ e S^2 são v.a. independentes

1.3.2 - Distribuição t de Student

Sejam $Z \sim N(0,1)$, $U \sim \chi^2_n$, Z e U v.a. independentes. Nestas condições, a v.a.

$$T = \frac{Z}{\sqrt{\frac{U}{n}}}$$

tem distribuição t-student com n graus de liberdade e sua função densidade de probabilidades é

$$f(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} \quad -\infty < t < \infty$$

Notação: $T \sim t_n$

Teorema

Se $Y_1, Y_2 \ldots Y_n$ é uma amostra aleatória da distribuição $N(\mu, \sigma^2)$, $\bar{Y}$ e $\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \bar{Y})^2$ são a média e a variância amostral, então

$$T = \frac{(\bar{Y} - \mu)}{\hat{\sigma}/\sqrt{n}} \sim t_{n-1}$$

Base do teste t para a média com variância desconhecida

Obs:

A função geradora de momentos da v. a. com distribuição t-student não existe

Para $n > 1$, $E(T) = 0$

Para $n > 2$, $Var(T) = \frac{n}{n-2}$

1.3.3 - Distribuição F de Snedecor

Se U_1 e U_2 são v.a. independentes, $U_1 \sim \chi^2_{n_1}$ e $U_2 \sim \chi^2_{n_2}$, então

$$W = \frac{\frac{U_1}{n_1}}{\frac{U_2}{n_2}}$$ tem dist. F de Snedecor com n_1 g.l. no numerador e n_2 no denominador. A função densidade de probabilidades de W é dada por

$$f(w) = \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)\left(\frac{n_1}{n_2}\right)^{n_1/2}}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)} w^{\frac{n_1-2}{2}}\left(1+\frac{n_1}{n_2}w\right)^{-\frac{n_1+n_2}{2}}$$

$w > 0$

Notação: $W \sim F_{n_1, n_2}$

Resultados

- A função geradora de momentos de W não existe
- Se $T \sim t_n$ então $T^2 \sim F_{1,n}$
- Se $W \sim F_{n_1;n_2}$ então $E(W) = \dfrac{n_2}{n_2-2}$ $\quad n_2 > 2$

$$Var(W) = \frac{2n_2^2(n_1+n_2-2)}{n_1(n_2-2)^2(n_2-4)} \qquad E(W^2) = \frac{n_2^2(n_1+2)}{n_1(n_2-2)(n_2-4)} \quad n_2 > 4$$

1.3.4- Distribuição Qui-Quadrado não Central

$$X_i \sim N(0,1) \Rightarrow \sum_{i=1}^{n} X_i^2 \sim \chi_n^2 \qquad X_i \sim N(\mu, \sigma^2) \Rightarrow \sum_{i=1}^{n} \left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi_n^2$$

$$X_i \sim N(\mu, 1) \Rightarrow \sum_{i=1}^{n} X_i^2 \sim \chi_n^2 \text{ não central } \quad \lambda = \frac{n\mu^2}{2}$$

$$X_i \sim N(\mu, \sigma^2) \Rightarrow \sum_{i=1}^{n} \frac{X_i^2}{\sigma^2} \sim \chi_n^2 \text{ não central } \quad \lambda = \frac{n\mu^2}{2\sigma^2}$$

A v.a. U tem distribuição qui-quadrado não central com n (inteiro positivo) graus de liberdade e parâmetro de não centralidade λ se

$$f(u) = e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \frac{u^{(n+2j-2)/2} \, e^{-u/2}}{\Gamma\left(\frac{n+2j}{2}\right) 2^{j+\frac{n}{2}}} \qquad u > 0$$

Obs:

Define-se $\lambda^j = 1$ para $\lambda = 0$ e $j = 0$. Para $\lambda = 0$, esta densidade se reduz à da v.a. com dist. χ^2_n

Resultados

1- $m_U(t) = (1-2t)^{-n/2} \exp\left(\frac{2t\lambda}{1-2t}\right), \quad t < \frac{1}{2}$

2- $E(U) = n + 2\lambda$

$Var(U) = 2(n + 4\lambda)$

3- Se U_1 e U_2 são v.a. independentes, $U_1 \sim \chi^2_{n_1, \lambda_1}$ e $U_2 \sim \chi^2_{n_2, \lambda_2}$, então $U_1 + U_2 \sim \chi^2_{n_1+n_2, \lambda_1+\lambda_2}$

Teorema

Sejam $X_i \sim N(\mu_i, \sigma^2)$ $i = 1, 2 \dots n$ variáveis aleatórias independentes. Nestas condições,

$$\sum_{i=1}^{n} \frac{X_i^2}{\sigma^2} \sim \chi^2_{n, \lambda}$$

onde $\lambda = \dfrac{\sum_{i=1}^{n} \mu_i^2}{2\sigma^2}$

→ Veremos posteriormente com o enfoque de formas quadráticas

1.3.5 - Distribuição F não central

Se $U_1 \sim \chi^2_{n_1, \lambda}$ e $U_2 \sim \chi^2_{n_2}$, U_1 e U_2 independentes

e $W = \dfrac{\frac{U_1}{n_1}}{\frac{U_2}{n_2}}$ tem distribuição F não central com n_1

graus de liberdade no numerador, n_2 graus de liberdade no denominador e parâmetro de não centralidade λ, cuja densidade é

$$f(w) = \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} \frac{n_1^{\frac{n_1}{2}+k}\, n_2^{\frac{n_2}{2}}\, \Gamma\left(\frac{n_1}{2}+\frac{n_2}{2}+k\right)}{\Gamma\left(\frac{n_1}{2}+k\right)\Gamma\left(\frac{n_2}{2}\right)} \frac{w^{\frac{n_1}{2}+k-1}}{(n_2+n_1 w)^{\frac{n_1}{2}+\frac{n_2}{2}+k}}$$

Notação: $W \sim F_{n_1, n_2, \lambda}$ ($\lambda = 0$, $f(w)$ se reduz à densidade F_{n_1, n_2})

Resultados

i) $$E(W) = \frac{n_2}{n_2 - 2}\left(1 + \frac{2\lambda}{n_1}\right)$$

$$Var(W) = \frac{2n_2^2}{n_1^2(n_2-2)}\left[\frac{(n_1+2\lambda)^2}{(n_2-2)(n_2-4)} + \frac{n_1+4\lambda}{n_2-4}\right]$$

Aplicações

Modelo de Regressão Linear Simples

$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ $\quad \epsilon_i \sim N(0, \sigma^2)$ independentes $i = 1, 2 \dots n$

$\hat{\beta}_0, \hat{\beta}_1$ estimadores de mínimos quadrados de β_0 e β_1

SSR - soma de quadrados da regressão $= \sum (\hat{Y}_i - \bar{Y})^2$

SSE - soma de quadrados do resíduo $= \sum (Y_i - \hat{Y}_i)^2$

1 - $\dfrac{SSE}{\sigma^2} \sim \chi^2_{n-2}$ $\qquad SSE \sim \text{prop } \chi^2_{n-2}$

2 - $\dfrac{SSR}{\sigma^2} \sim \chi^2_{1,\lambda}$ $\qquad \lambda = \dfrac{\beta_1^2 \sum (x_i - \bar{x})^2}{2\sigma^2}$

e são independentes

Se $\beta_1 = 0$ $\quad \dfrac{SSR}{\sigma^2} \sim \chi^2_1$ e $F = \dfrac{\frac{SSR}{\sigma^2}}{\frac{SSE}{(n-2)\sigma^2}} = \dfrac{SSR}{\frac{SSE}{n-2}} \sim F_{1,n-2}$

(central)

Rejeitamos $H_0: \beta_1 = 0$ para $\overbrace{\dfrac{SSR}{\frac{SSE}{n-2}}}^{F} \geq F_c$

$F_c \mid \quad P(F_{1,n-2} \geq F_c) = \alpha$

Cálculo do Poder do Teste

$$\Pi(b) = P(\text{rejeitar } H_0 \mid \beta_1 = b)$$

$$= P(F \geq F_c \mid \beta_1 = b) = \int_{F_c}^{+\infty} F(w, 1, n-2, \lambda)\, dw$$

onde $F(w, 1, n-2, \lambda)$ é a densidade de uma v.a.

com distribuição $F_{1, n-2, \lambda}$ e $\lambda = \dfrac{b^2 \sum (x_i - \bar{x})^2}{2\sigma^2}$

Ver Tabela T.11 - Graybill

substituir σ^2 por $\hat{\sigma}^2 = \dfrac{SSE}{n-2}$

Quanto à $\int_{F_c}^{\infty} F(w, n_1, n_2, \lambda)\, dw$:

a) é monotônica crescente em λ, fixados n_1, n_2 e α

Qto mais distante β_1 de zero, maior o poder

b) é monotônica crescente em α, fixados n_1, n_2 e λ

$> \alpha$, $< F_c$, maior a cauda

c) é monotônica crescente em n_2 fixados α, n_1 e λ

(Qto maior o tamanho da amostra, maior o poder)

fixado o valor da alternativa b

d) é monotônica decrescente em n_1 fixados n_2, α e λ.

Qto mais v. independentes, a situação piora

Vire →

1.3.6 – Outras distribuições não centrais

Distribuição t de student não central

Se $X \sim N(\mu, 1)$ e $U \sim \chi^2_n$, X e U v. a. independentes, então

$$T^* = \frac{X}{\sqrt{\frac{U}{n}}}$$

aplicações:
Cálculo do poder para o teste de média da normal, Var. desconhecida

tem distribuição t de student não central, com n graus de liberdade e parâmetro de não centralidade μ.

Sua densidade é

$$f(t) = \frac{n^{\frac{1}{2}n}}{\Gamma\left(\frac{n}{2}\right)} \frac{e^{-\frac{1}{2}\mu^2}}{(n+t^2)^{\frac{1}{2}(n+1)}} \sum_{k=0}^{\infty} \frac{\Gamma\left(\frac{1}{2}n + \frac{1}{2}k + \frac{1}{2}\right)\mu^k 2^{\frac{1}{2}k} t}{k!\,(n+t^2)^{\frac{1}{2}k}}$$

Tabelada, Tabela 4.3.1 - Graybill

Distribuição F não central dupla

$U_1 \sim \chi^2_{n_1, \lambda_1}$ $\quad$ $U_2 \sim \chi^2_{n_2, \lambda_2}$ independentes

$$F^* = \frac{\frac{U_1}{n_1}}{\frac{U_2}{n_2}} \sim F_{n_1, n_2, \lambda_1, \lambda_2}$$

Densidade: pag 53 Searle