9.7 - De Groot

(1) Deriva a função de densidade de probabilidade de uma distribuição F (#9.7.1)

Seja as variáveis aleatórias Y e W independentes com $Y \sim \chi^2(m)$ e $W \sim \chi^2(n)$, sendo conjunta $g(y,w)$ é o produto das p.d.f individuais. Além disso como Y e W tem mesma distribuição e é χ^2, é dado em 8.2.1. ~~que~~ ~~g~~ o valor de $g(y,w)$ com $y>0$ e $w>0$,

$$g(y,w) = c\, y^{(m/2)-1} w^{(n/2)-1} e^{-(y+w)/2}$$

onde,
$$c = \frac{1}{2^{(m+n)/2}\,\Gamma\left(\frac{1}{2}m\right)\Gamma\left(\frac{1}{2}n\right)}$$

Seja, $X = \frac{Y/m}{W/n}$, a função p.d.f conjunta $h(x,w)$ de X e W é obtida substituindo primeiro Y, temos $y = \left(\frac{m}{n}\right)xw$ e $\frac{\partial y}{\partial x} = \left(\frac{m}{n}\right)w$. Assim podemos encontrar a forma de $x>0$ e $w>0$.

$$h(x,w) = c\left(\frac{m}{n}\right)^{m/2} x^{(m/2)-1}\, w^{[(m+n)/2]-1} \exp\left[-\frac{1}{2}\left(\frac{m}{n}x+1\right)w\right]$$

A ~~margem~~ pdf marginal $f(x)$ de X pode ser obtida para cada valor de $x>0$ ~~com a~~ para a relação

$$f(x) = \int_0^\infty h(x,w)\,dw$$

E segue que

$$\int_0^\infty w^{[(m+n)/2-1]} \exp\left[-\frac{1}{2}\left(\frac{m}{n}x+1\right)w\right] dw = \frac{\Gamma\left(\frac{1}{2}(m+n)\right)}{\left[\frac{1}{2}\left(\frac{m}{n}x+1\right)\right]^{\frac{m+n}{2}}}$$

② Discuta o teste F com um teste de razão de verossimilhança.

Supondo que observamos $X_1, \dots, X_m$ e $Y_1, \dots, Y_n$ sendo duas amostras, a função de verossimilhança $g(x,y|\mu_1,\mu_2,\sigma_1^2,\sigma_2^2)$ é

$$g(x,y|\mu_1,\mu_2,\sigma_1^2,\sigma_2^2) = f_m(x|\mu_1,\sigma_1^2)\, f_n(y|\mu_2,\sigma_2^2)$$

Aqui, tanto f_m e f_n tem a forma geral

$$f_n = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[\frac{1}{2\sigma^2}\sum (x_i-\mu)^2\right]$$

~~usando~~ Para as hipóteses em (9.7.3), Ω_0 contem todos os parametros $\theta = (\mu_1,\mu_2,\sigma_1^2,\sigma_2^2)$ com $\sigma_1^2 \le \sigma_2^2$, e Ω_1 contem todos todos os θ com $\sigma_1^2 > \sigma_2^2$. A ~~verossimilhança~~ estatística de razão de verossimilhança é

$$\Lambda(x,y) = \frac{\sup_{\{(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2):\ \sigma_1^2 \le \sigma_2^2\}} g(x,y|\mu_1,\mu_2,\sigma_1^2,\sigma_2^2)}{\sup_{\{\mu_1,\mu_2,\sigma_1^2,\sigma_2^2\}} g(x,y|\mu_1,\mu_2,\sigma_1^2,\sigma_2^2)}$$

O teste de razão de verossimilhança especifica que H_0 deve ser rejeitado se $\Lambda(x,y) \le k$, onde k é tipicamente escolhido para fazer o teste de nível desejado.

Para facilitar as maximizações, seja

$$S_x^2 = \sum^m (x_i - \bar{x}_m)^2, \quad e \quad S_y^2 = \sum^n (y_i - \bar{y}_n)^2$$

Então, podemos escrever

$$g(x,y|\mu_1,\mu_2,\sigma_1^2,\sigma_2^2) = \frac{1}{(2\pi)^{(m+n)/2}\sigma_1^m \sigma_2^n} \exp\left(-\frac{1}{2\sigma_1^2}\left[m(\bar{x}_m-\mu_1)^2 + S_x^2\right] - \frac{1}{2\sigma_2^2}\left[n(\bar{y}_n-\mu_2)^2 + S_y^2\right]\right)$$

Tanto para o numerador quanto o denominador precisamos que $\mu_1 = \bar{x}_m$ e $\mu_2 = \bar{y}_n$ para maximizar a verossimilhança.
Se $S_x^2/m \le S_y^2/n$, então o numerador é maximizado em $\sigma_1^2 = S_x^2/m$ e $\sigma_2^2 = S_y^2/n$. Esses valores também maximizam o denominador.

Portanto, $\Lambda(x,y)=1$ se $S_x^2/m \leq S_y^2/n$. Para o caso contrário, podemos mostrar que o máximo é atingido em $\sigma_1^2=\sigma_2^2$ onde

$$\sigma_1^2=\sigma_2^2=\frac{S_x^2+S_y^2}{m+n}$$

Substituindo, temos

$$\Lambda(x,y)=\begin{cases} 1, & S_x^2/m \leq S_y^2/n \\ d\,w^{m/2}(1-w)^{n/2}, & \text{caso contrário} \end{cases}$$

onde,

$$w=\frac{S_x^2}{S_x^2+S_y^2} \quad \text{e} \quad d=\frac{(m+n)^{(m+n)/2}}{m^{m/2}\,n^{n/2}}$$