ME720 - Modelos Lineares Generalizados

Parte 5 - Poder do teste e determinação do tamanho da amostra em modelos normais lineares

Profa. **Larissa Avila Matos**

Poder do teste

Sabemos que,

- α = probabilidade do erro do tipo I = P(Rejeitar $H_0 | H_0$ é verdadeira);
- β = probabilidade do erro do tipo II = P(Não rejeitar $H_0 | H_0$ é falsa);
- ψ = poder do teste = $1 - \beta$ = P(Rejeitar $H_0 | H_0$ é falsa).

Em geral, em qualquer experimento, a probabilidade do erro do tipo I é controlada (α).

A probabilidade do erro do tipo II (consequentemente o poder do teste) não é, em geral, controlada.

Como determinar tamanhos de amostra (por tratamento/no geral) que garantam um poder mínimo?

Como calcular o poder do teste, para um dado experimento?

Distribuição qui-quadrado não central

Sejam $X_1, X_2, ..., X_n \overset{ind}{\sim} N(\mu_i, \sigma_i^2)$.

Defina $Y = \sum_{i=1}^{n} \left(\frac{X_i}{\sigma_i}\right)^2$. Dizemos então que Y tem distribuição qui-quadrado não central com n graus de liberdade e parâmetro de não-centralidade $\delta = \sum_{i=1}^{n} \left(\frac{\mu_i}{\sigma_i}\right)^2$.

Notação: $Y \sim \chi^2_{(n,\delta)}$, cuja fdp é dada por

$$f_y(y) = \sum_{i=0}^{\infty} \frac{e^{-\delta/2}(\delta/2)^i}{i!} f_{W_{n+2i}}(y) \mathbb{1}_{(0,\infty)}(y),$$

em que $W_{n+2i} \sim \chi^2_{(n+2i)}$.

Se $\delta = 0$, então $Y \sim \chi^2_{(n)}$.

Distribuição F não central

Seja V uma outra v.a., independente de Y, $V \sim \chi^2_{(m)}$.

Defina $F = \frac{Y/n}{V/m}$. Então, F tem distribuição F não central com graus de liberdade, n e m e parâmetro de não centralidade δ.

Notação: $F \sim F_{(n,m,\delta)}$, cuja fdp é dada por

$$\begin{aligned} f_F(f) &= \sum_{i=0}^{\infty} \frac{e^{-\delta/2}(\delta/2)^i}{\beta(m/2, n/2+i)i!} \left(\frac{n}{m}\right)^{n/2+i} \left(\frac{m}{m+nf}\right)^{(n+m)/2+i} \\ &\times f^{n/2-1+i} \mathbb{1}_{(0,\infty)}(f), \end{aligned}$$

em que $\boldsymbol{\beta}(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx$.

Modelo para um único fator

As observações podem ser descritas pelo modelo,

$$Y_{ij} = \mu + \alpha_i + \xi_{ij},$$

com $i = 1, 2, .., k$ (grupos); $j = 1, ..., n$ (u.e).

- Erros $\xi_{ij} \overset{i.i.d}{\sim} N(0, \sigma^2)$.
- μ, α_i não aleatórios.
- Restrição: $\alpha_1 = 0$ (Casela de Referência).

Hipótese de interesse (primário)

$$\begin{aligned} H_0 : &\quad \alpha_2 = \alpha_3 = ... = \alpha_k = 0 \quad vs \\ H_1 : &\quad \text{pelo menos uma desigualdade.} \end{aligned}$$

Lembremos que, sob $H_0, V = SQF/\sigma^2 \sim \chi^2_{(k-1)}$.

Independentemente de H_0 ser verdadeira, $W = SQR/\sigma^2 \sim \chi^2_{(n-k)}$.

Assim, sob H_0 $F = \frac{V/(k-1)}{W/(n-k)} \sim F_{(k-1,n-k)}$.

Lembrando que,

$$\mathbb{E}(SQF) = (k-1)\sigma^2 + n\sum_{i=1}^{k}(\mu_i - \overline{\mu})^2, \quad \text{com} \quad \overline{\mu} = \frac{1}{k}\sum_{i=1}^{k}\mu_i$$

No caso da parametrização casela de referência, tem-se que $\mu_i - \overline{\mu} = \alpha_i - \overline{\alpha}$, $\overline{\alpha} = \frac{1}{k}\sum_{i=1}^{k}\alpha_i$.

Assim,

$$\mathbb{E}(SQF/\sigma^2) = \underbrace{(k-1)}_{\text{graus de liberdade}} + \underbrace{\frac{n}{\sigma^2}\sum_{i=1}^{k}(\alpha_i - \overline{\alpha})^2}_{\text{parâmetro de não centralidade}}.$$

Assim, se H_0 não for verdadeira, então

$$V \sim \chi^2_{(k-1,\delta)}, \delta = n \frac{\sum_{i=1}^{k} (\alpha_i - \overline{\alpha})^2}{\sigma^2}.$$

Portanto,

$$\mathbb{E}(SQF/\sigma^2) = (k-1) + \delta.$$

Segue-se que, sob H_1, $F \sim F_{(k-1, nk-k, \delta)}$.

Regra para obtenção do parâmetro de não-centralidade. Em geral,

$$\mathbb{E}(\text{Soma de Quadrados Fator}/\sigma^2) = \text{g.l.} + \text{parâmetro de não centralidade}$$

Assim, para um dado valor de δ e para um nível de significância fixado α, temos que o poder do teste é dado por

$$\psi = P(F_1 > f_c | H_0 \text{ é falsa}), F_1 \sim F_{(k-1, nk-k, \delta)},$$

em que f_c (valor crítico) é o valor da distribuição $F_{(k-1,nk-k)}$, $P(F_0 > f_c | H_0) = \alpha$, $(F_0 \sim F_{(k-1,nk-k)})$.

A função poder (em δ) é dada por:

$$\psi(\delta) = P(F_1 > f_c | H_0 \text{ é falsa}), \quad F_1 \sim F_{(k-1,nk-k,\delta)} \tag{1}$$

Para um determinado conjunto de dados, podemos estimar δ, através de

$$\widehat{\delta} = n\frac{\sum_{i=1}^{k}(\widehat{\alpha}_i - \widehat{\overline{\alpha}})^2}{\widehat{\sigma}^2},$$

em que $\widehat{\alpha}_i, i = 1, 2.., k$ são os estimadores de mínimos quadrados,
$\widehat{\overline{\alpha}} = \frac{1}{k}\sum_{i=1}^{k}\widehat{\alpha}_i$ e $\widehat{\sigma}^2 = QMR$.

Logo, o poder estimado é dado por

$$\widehat{\psi} = \psi(\widehat{\delta}),$$

ou seja, utiliza-se $\widehat{\delta}$ na equação (1).

Exemplo Solvente

Uma bioquímica (Tecnologia de Alimentos) está interessada em estudar a extração de pigmentos naturais, com aplicação como corante em alimentos. Numa primeira etapa tem-se a necessidade de escolher o melhor solvente extrator. A escolha do(s) melhor(es) solventes foi realizada através da medida da absorbância de um pigmento natural do fruto de baguaçú.

- Fator = tipos de solvente;
- k=5 níveis;
- n_k=5 repetições.

```r
fit.model <- lm(mabsor~solvfac)
summary(fit.model)
```

```
Call:
lm(formula = mabsor ~ solvfac)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.04024 -0.01874 -0.00086  0.01914  0.05136 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.53934    0.01128  47.826  < 2e-16 ***
solvfacE70   0.06854    0.01595   4.298 0.000351 ***
solvfacEAW   0.02752    0.01595   1.726 0.099843 .  
solvfacM1M  -0.34258    0.01595 -21.481 2.75e-15 ***
solvfacMAW  -0.08970    0.01595  -5.624 1.67e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02522 on 20 degrees of freedom
Multiple R-squared:  0.977, Adjusted R-squared:  0.9725 
F-statistic: 212.8 on 4 and 20 DF,  p-value: 4.378e-16
```

```r
fit.model$coef
```

```
(Intercept)  solvfacE70  solvfacEAW  solvfacM1M  solvfacMAW
    0.53934     0.06854     0.02752    -0.34258    -0.08970
```

```r
sigma2 <- (summary(fit.model)$sigma)^2
sigma2
```

```
[1] 0.0006358708
```

```r
sum(fit.model$residuals^2)/fit.model$df.residual
```

```
[1] 0.0006358708
```

```r
v.alpha <- c(0,fit.model$coef[-1])
delta <- 5*sum((v.alpha - mean(v.alpha))^2)/sigma2
delta
```

```
[1] 851.2221
```

Da análise anterior, temos:

- $\widetilde{\boldsymbol{\alpha}} = (0;\ 0,06854;\ 0,02752;\ -0,34258;\ -0,08970)$; e
- $\widetilde{\sigma}^2 = 0,0006358708$.

Isto implica que $\widetilde{\delta} = 851,222$.

Neste caso, para um $\alpha = 0,05$, o valor crítico é igual à $2,866$.

```
qf(p=0.95,df1=5-1,df2=25-5)
```

```
[1] 2.866081
```

Assim, o poder estimado será igual à

$$\widetilde{\psi} = P(F_1 > 2,866 | H_0 \text{ é falsa}) \equiv P(F_1 > 2,866) > 0,9999.$$

```
pf(q=2.866081,df1=5-1,df2=25-5,ncp=delta)
```

```
[1] 1.440173e-123
```

Exemplo

Tem-se o interesse em se saber se a quantidade de fósforo existente (administrada) no solo afeta a produção de milho (de uma certa variedade).

Fator: quantidade de fósforo, k = 5 níves, $n = 4, i = 1, 2, 3, 4$ repetições por tratamento (quantidade de fósforo administrada).

Procedimento: 20 porções de terras, chamadas de parcelas (em condições semelhantes) foram consideradas e cada uma delas recebeu uma determinada quantidade de fósforo, de modo aleatório (completamente casualizado).

Resposta: produção de milho (kg/parcela).

Experimento balanceado.

Dados:

Quantidade de Fósforo	Produção			
	1	2	3	4
0 kg/ha	2,38	6,77	3,50	5,94
25 kg/ha	6,15	8,78	8,99	9,10
50 kg/ha	9,07	8,73	6,92	8,48
75 kg/ha	9,55	8,95	10,24	8,66
100 kg/ha	9,14	10,17	9,75	9,50

Modelo Casela de Referência

As observações podem ser descritas pelo modelo:

$$Y_{ij} = \mu + \alpha_i + \xi_{ij},$$

com $i = 1, 2, .., 5$ (tratamentos); $j = 1, ..., 4$ (u.e).

- Erros $\xi_{ij} \overset{i.i.d}{\sim} N(0, \sigma^2)$.
- μ, α_i não aleatórios.
- Restrição: $\alpha_1 = 0$.

```r
fit.model <- lm(pmilho~fosfac)
summary(fit.model)
```

```

Call:
lm(formula = pmilho ~ fosfac)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.2675 -0.5475  0.1900  0.7438  2.1225 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   4.6475     0.6237   7.451 2.04e-06 ***
fosfac25      3.6075     0.8820   4.090 0.000966 ***
fosfac50      3.6525     0.8820   4.141 0.000871 ***
fosfac75      4.7025     0.8820   5.331 8.39e-05 ***
fosfac100     4.9925     0.8820   5.660 4.53e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.247 on 15 degrees of freedom
Multiple R-squared:  0.7315,	Adjusted R-squared:  0.6599 
F-statistic: 10.22 on 4 and 15 DF,  p-value: 0.0003379
```

```r
anova(fit.model)
```

```
Analysis of Variance Table

Response: pmilho
          Df Sum Sq Mean Sq F value    Pr(>F)    
fosfac     4 63.596  15.899  10.218 0.0003379 ***
Residuals 15 23.340   1.556                      
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Estimativas dos parâmetros do modelo

Parâmetro	Estimativa	EP	IC(95%)	Estat. t	p-valor
μ	4,6475	0,6237	[3,318; 5,977]	7,451	<0,001
α_2	3,6075	0,8820	[1,727; 5,488]	4,090	<0,001
α_3	3,6525	0,8820	[1,772; 5,533]	4,141	<0,001
α_4	4,7025	0,8820	[2,822; 6,583]	5,331	<0,001
α_5	4,9925	0,8820	[3,112; 6,873]	5,660	<0,001

Estimativas das médias

Parâmetro	Estimativa	EP	IC(95%)
μ_1	4,647	0,624	[3,318; 5,977]
μ_2	8,255	0,624	[6,926; 9,584]
μ_3	8,300	0,624	[6,971; 9,629]
μ_4	9,350	0,624	[8,021; 10,679]
μ_5	9,640	0,624	[8,311; 10,969]

```r
fit.model$coef
```

```
(Intercept)    fosfac25    fosfac50    fosfac75   fosfac100
     4.6475      3.6075      3.6525      4.7025      4.9925
```

```r
sigma2 <- (summary(fit.model)$sigma)^2
sigma2
```

```
[1] 1.556012
```

```r
sum(fit.model$residuals^2)/fit.model$df.residual
```

```
[1] 1.556012
```

```r
v.alpha <- c(0,fit.model$coef[-1])
delta <- 4*sum((v.alpha - mean(v.alpha))^2)/sigma2
delta
```

```
[1] 40.87108
```

Da análise anterior, temos:

- $\widetilde{\boldsymbol{\alpha}} = (0;\ 3,6075;\ 3,6525;\ 4,7025;\ 4,9925)$ e
- $\widetilde{\sigma}^2 = 1,556012$;

Logo $\widetilde{\delta} = 40,871$.

Assim, temos os seguintes resultados, conforme o valor de α escolhido.

α	f_c	$\widetilde{\psi}$
0,10	2,36	0,9992
0,05	3,06	0,9965
0,01	4,89	0,9624

```r
qf(0.90,5-1,20-5)
```

```
[1] 2.361433
```

```r
1-pf(qf(0.90,5-1,20-5),5-1,20-5,delta)
```

```
[1] 0.9992473
```

```r
qf(0.95,5-1,20-5)
```

```
[1] 3.055568
```

```r
1-pf(qf(0.95,5-1,20-5),5-1,20-5,delta)
```

```
[1] 0.9965949
```

```r
qf(0.99,5-1,20-5)
```

```
[1] 4.89321
```

```r
1-pf(qf(0.99,5-1,20-5),5-1,20-5,delta)
```

```
[1] 0.9624809
```

Suponha que o pesquisador queira realizar um outro experimento semelhante à este em questão.

Ele quer, para $\alpha = 0,05$, um poder de pelo menos 0,90.

Qual deve ser o tamanho para cada tratamento (consequentemente o tamanho total) para obter este poder?

n	nk	f_c	δ	$\widetilde{\psi}$
2	10	5,19	20,43	0,6289
3	15	3,47	30,65	0,9536
4	20	3,06	40,87	0,9965
5	25	2,87	51,09	0,9998

```r
k <- 5
ni <- c(2,3,4,5)
ptests <- NULL
deltas <- NULL
quantfs <- NULL
for(i in 1:length(ni)){
n <- ni[i]
delta <- n*sum((v.alpha-mean(v.alpha))^2)/sigma2
deltas[i] <- delta
quantf <- qf(0.95,k-1,k*n-k)
quantfs[i] <- quantf
ptests[i] <- 1-pf(quantf,k-1,k*n-k,delta)}
cbind(n=ni,nk=ni*k,fc=round(quantfs,2),delta=round(deltas,2),
      poder=round(ptests,4))
```

```
     n nk   fc delta  poder
[1,] 2 10 5.19 20.44 0.6289
[2,] 3 15 3.48 30.65 0.9536
[3,] 4 20 3.06 40.87 0.9966
[4,] 5 25 2.87 51.09 0.9998
```

Como proceder

Perguntar ao pesquisador e/ou usar informações de experimentos anteriores sobre o nível de significância (α).

Perguntar ao pesquisador e/ou usar informações de experimentos anteriores sobre a variabilidade (σ^2).

Perguntar ao pesquisador e/ou usar informações de experimentos anteriores sobre a magnitude das diferenças entre as médias.

Exemplo artificial

Três grupos, experimento balanceado, $\alpha = 0,05$, $\sigma^2 = 5$, $k = 3$.

			$\sum_{i=1}^{3} \alpha_i^2$							
			0, 5		2		50		200	
n	kn	f_c	δ	$\widetilde{\psi}$	δ	$\widetilde{\psi}$	δ	$\widetilde{\psi}$	δ	$\widetilde{\psi}$
2	6	9,55	0,07	0,0522	0,30	0,0599	6,80	0,2790	30,24	0,7907
3	9	5,14	0,10	0,0549	0,45	0,0721	10,20	0,5882	45,36	0,9963
4	12	4,26	0,14	0,0575	0,60	0,0846	13,60	0,7970	60,48	>0,9999
5	15	3,89	0,17	0,0602	0,76	0,0974	17,00	0,9091	75,60	>0,9999
10	30	3,35	0,34	0,0737	1,51	0,1649	34,01	0,9993	151,20	>0,9999

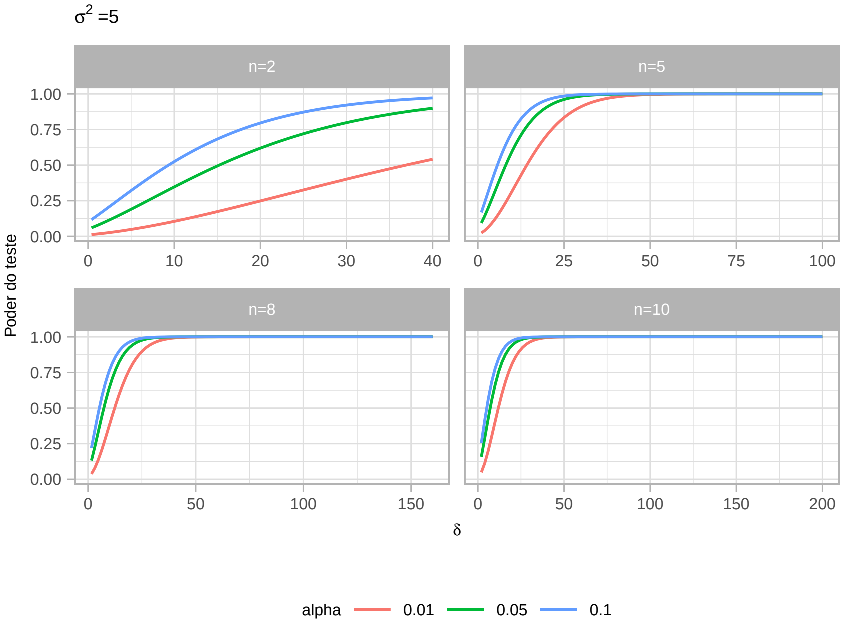
```
k <- 3
ni <- c(2,3,4,5,10)
alphas<-rbind(sqrt(c(0,0.2,0.3)),sqrt(c(0,0.5,1.5)),sqrt(c(0,20,30)),sqrt(c(0,50,150)))
sigma2<-5
ptests <- NULL
deltas <- NULL
quantfs <- NULL
x=NULL
for(j in 1:nrow(alphas)){
  v.alpha<-alphas[j,]
for(i in 1:length(ni)){
n <- ni[i]
delta <- n*sum((v.alpha-mean(v.alpha))^2)/sigma2
deltas[i] <- delta
quantf <- qf(0.95,k-1,k*n-k)
quantfs[i] <- quantf
ptests[i] <- 1-pf(quantf,k-1,k*n-k,delta)}
x=cbind(x,n=ni,nk=ni*k,fc=round(quantfs,2),delta=round(deltas,2),poder=round(ptests,4))}
x
```

```
      n nk   fc delta  poder  n nk   fc delta  poder  n nk   fc delta
[1,]  2  6 9.55  0.07 0.0522  2  6 9.55  0.30 0.0599  2  6 9.55  6.80
[2,]  3  9 5.14  0.10 0.0549  3  9 5.14  0.45 0.0721  3  9 5.14 10.20
[3,]  4 12 4.26  0.14 0.0575  4 12 4.26  0.60 0.0846  4 12 4.26 13.60
[4,]  5 15 3.89  0.17 0.0602  5 15 3.89  0.76 0.0974  5 15 3.89 17.00
[5,] 10 30 3.35  0.34 0.0737 10 30 3.35  1.51 0.1649 10 30 3.35 34.01
      poder  n nk   fc  delta  poder
[1,] 0.2790  2  6 9.55  30.24 0.7907
[2,] 0.5882  3  9 5.14  45.36 0.9963
[3,] 0.7970  4 12 4.26  60.48 1.0000
[4,] 0.9091  5 15 3.89  75.60 1.0000
[5,] 0.9993 10 30 3.35 151.20 1.0000
```

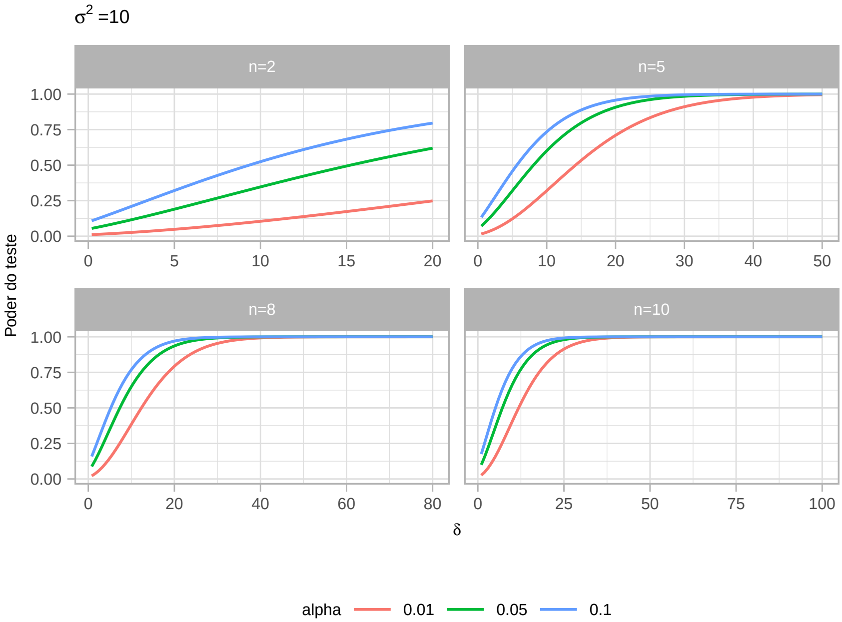
Curvas do poder do teste

Configurações:

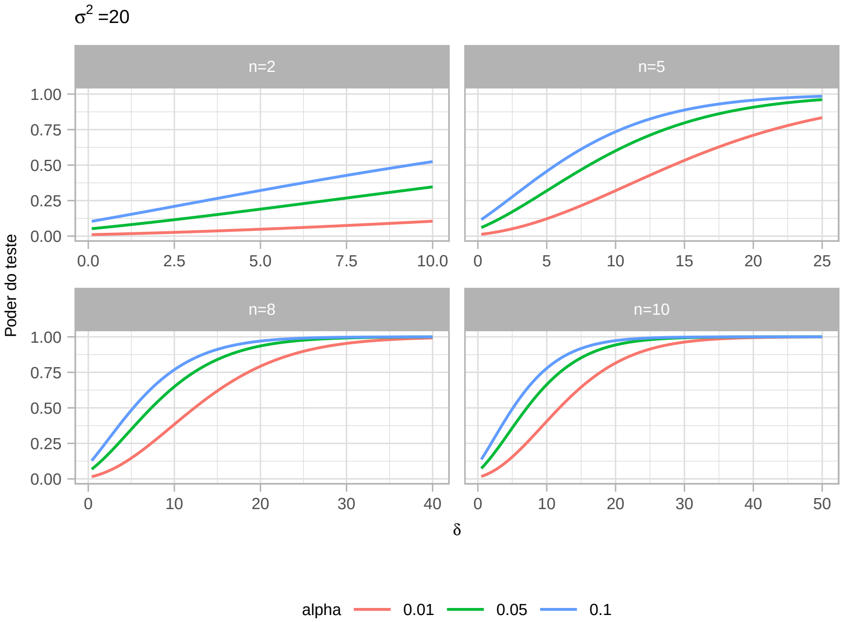
- $\sigma^2 = 5,\ 10$ e 20
- $n = 2,\ 5,\ 8$ e 10
- $\alpha = 0,01,\ 0,05,$ e $0,10$



σ² =5
n=2
n=5
n=8
n=10
1.00
0.75
0.50
0.25
0.00
0
10
20
30
40
0
25
50
75
100
0
50
100
150
0
50
100
150
200
Poder do teste
δ
alpha
0.01
0.05
0.1



σ² =10
n=2
n=5
n=8
n=10
Poder do teste
δ
alpha
0.01
0.05
0.1



σ^2 =20
n=2
n=5
n=8
n=10
Poder do teste
1.00
0.75
0.50
0.25
0.00
0.0
2.5
5.0
7.5
10.0
0
5
10
15
20
25
30
40
50
δ
alpha
0.01
0.05
0.1

Comentários sobre o poder do teste

- Podemos notar que, quanto maior o valor do parâmetro de não centralidade (δ), maior o poder do teste.

- Quanto maior for a diferença entre as médias e menor for a variância, maior será o valor de δ, consequentemente, maior será o poder do teste.

- Além disso, quanto maior for o valor da probabilidade do erro do tipo I (α), maior será o poder do teste;

Referência

- Notas de aula do Prof. Caio Azevedo.
- Agresti, A. (2015). *Foundations of Linear and Generalized Linear Models*. Wiley series in probability and statistics.