Trabalho 1: Teste Binomial

Estatística Não Paramétrica

Luiz Fernando Coelho Passos

02 de Setembro de 2019

1) Acredita-se que 30% das pessoas tenham problemas de audição. Numa amostra de 80 indivíduos 20 apresentaram o problema. Execute um teste binomial não paramétrico. Avalie o poder do teste.

Temos que, numa amostra de 80 indivíduos, 20 apresentaram o problema.

Pelo enunciado, acredita-se que 30% das pessoas tenham o problema. Assim, nossas hipóteses são:

$H_0 : p = 0.3$

$H_1 : p \neq 0.3$

Definindo T o número de pessoas que tem o problema, $T = 20$.

Sob a hipótese H_0, a variável aleatória T tem distribuição Binomial(80; 0.30)

Utilizaremos um nível de significância de 5%;

P-valor (Teste Binomial para grandes amostras)

Como T tem distribuição Binomial(80; 0.30), ou seja o tamanho da nossa amostra é 80, será utilizado a aproximação normal para a binomial, ou seja, assumiremos que T é assintoticamente normal com média np e variância $np(1-p)$. Assim,

$$P(Y \leq t_{obs}) \approx P(Z \leq \frac{t_{obs} - np^* + 0,5}{\sqrt{np^*(1-p^*)}})$$

e

$$P(Y \geq t_{obs}) \approx 1 - P(Z \leq \frac{t_{obs} - np^* + 0,5}{\sqrt{np^*(1-p^*)}})$$

onde, $t_{obs} = 20, n = 80$ e $p^* = 0.3$.

Assim, temos que,

```
# cálculos
# ( tobs-(np*)+0.5 )/ ( ( np*(1-p*) )^(1/2) )
( 20-(80*0.3)+0.5 )/( sqrt( (80*0.3)*(1-0.3) ) )
```

```
## [1] -0.8539126
```

$P(Y \leq 80) \approx P(Z \leq -0.85) = 1 - P(Z \leq 0.85) = 0,1977$

Logo, p-valor $= 2 \times P(Y \leq 80) = 0,3954$. Assim, com base no nível de significância de 5%, não rejeitamos H_0, ou seja, há evidências de que 30% das pessoas têm problemas de audição.

Calculando utilizando a distribuição binomial

Usando a distribuição binomial, o p-valor será obtido por:

$P(T \leq 20) + P(T \geq np^* + (np^* - 20)) = P(T \leq 20) + P(T \geq 28)$

```
# Cálculos
# (np*)+(np*-tobs)
(80*0.3)+(80*0.3-20)
```

```
## [1] 28
```

```
pbinom(20, 80, 0.3) + pbinom(28, 80, 0.3, lower.tail = F)
```

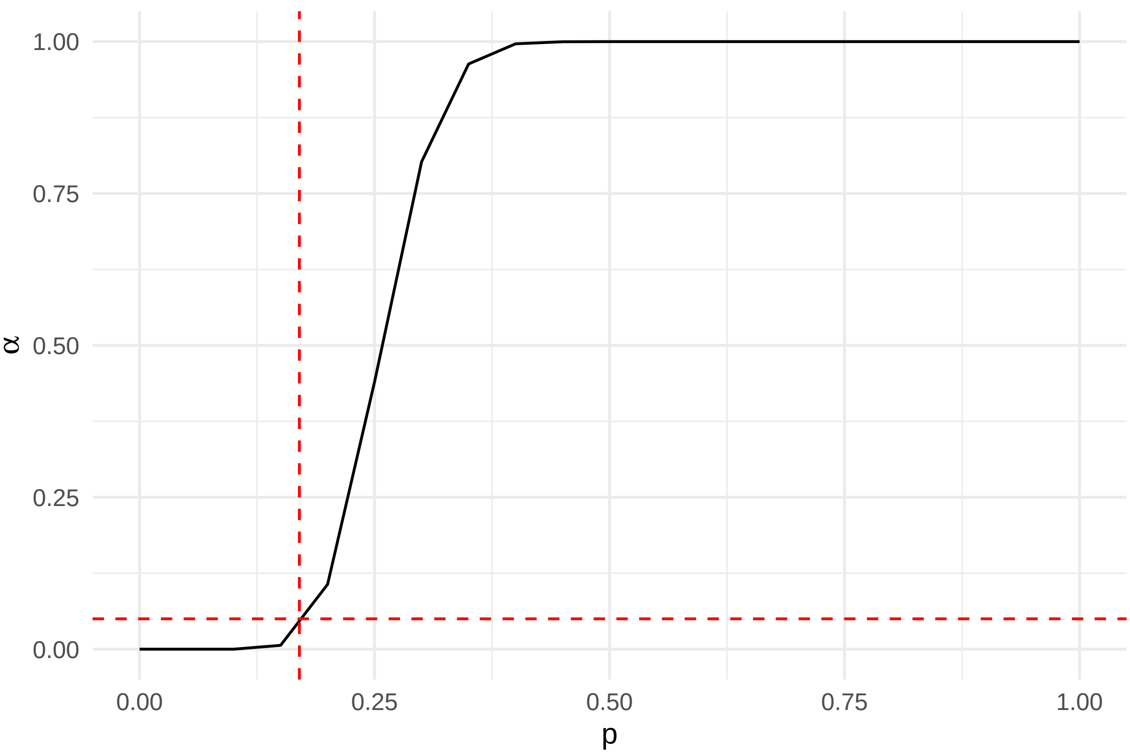
```
## [1] 0.3345324
```

$P(T \leq 20) + P(T \geq 28) = 0,3345$

Poder do teste

```
x = seq(0, 1, 0.05)
y = round( 1-pbinom(20,80,x), 4)

ggplot(data.frame(x), aes(x, y)) +
  geom_line() +
  geom_vline(xintercept = 0.17, color = "red", linetype = "dashed") +
  geom_hline(yintercept = 0.05, color = "red", linetype = "dashed") +
  labs(x = "p", y = expression(alpha)) +
  theme_minimal()
```



A partir de p=0,17, aproximadamente, o poder do teste é superior a 0,05.

2) Um industrial afirma que seu processo de fabricação produz 90% de peças dentro das especificações. Deseja-se investigar se este processo de fabricação ainda está sob controle. Uma amostra de 15 peças foi analisada e foram constatadas 10 peças dentro das especificações. Ao nível de 5% de significância, podemos dizer ser verdadeira essa afirmação?

Pelo enunciado temos que um industrial afirma que produz 90% de peças dentro das especificações. Temos também, uma amostra de 15 peças foi analisada e foram constatadas 10 peças dentro das especificações. Assim, nossas hipóteses são:

$H_0 : p = 0.9$

$H_1 : p \neq 0.9$

Definindo T como sendo o número de peças dentro das especificações, $T = 10$.

Sob a hipótese H_0, a variável aleatória T tem distribuição Binomial(15; 0.90)

Região Crítica e P-valor

Como nosso $\alpha = 0.05$ (nível de significância), queremos encontrar t_1 e t_2, tal que

$$P(T \leq t_1) = \alpha_1 \text{ e } P(T \geq t_2) = \alpha_2, \text{ onde } \alpha_1 + \alpha_2 = \alpha$$

Assim,

```
t1 = qbinom(0.0025, 15, 0.9)
t2 = qbinom(0.0025, 15, 0.9, lower.tail = F)
cat("t1 = ", t1, "\nt2 = ", t2)
```

```
## t1 =  10
## t2 =  15
```

Logo, o p-valor é igual a

```
pbinom(t1,15,0.9) + pbinom(t2,15,0.9, lower.tail = F)
```

```
## [1] 0.01272048
```

Teste Binomial

```
binom.test(10, 15, 0.9, alternative = "two.sided")
```

```
##
##  Exact binomial test
##
## data:  10 and 15
## number of successes = 10, number of trials = 15, p-value = 0.01272
## alternative hypothesis: true probability of success is not equal to 0.9
## 95 percent confidence interval:
##  0.3838037 0.8817589
## sample estimates:
## probability of success
##              0.6666667
```

O p-valor obtido com o teste binomial foi de 0,013. Logo, com base no nível de significância de 5%, rejeitamos H_0, ou seja, não há evidências de que o processo de fabricação produz 90% de peças dentro das especificações.