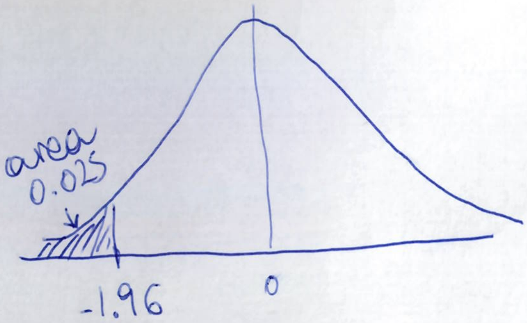
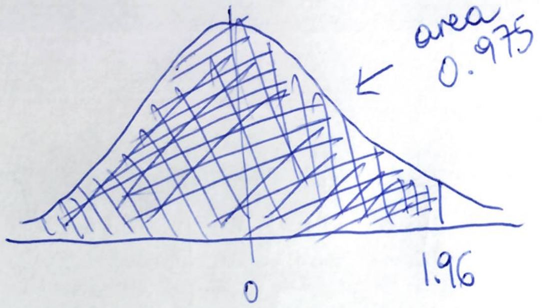
qnorm / pnorm

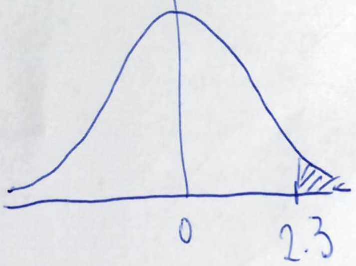


pnorm(-1.96) = 0.025
qnorm(0.025) = -1.96



pnorm(1.96) = 0.975
qnorm(0.975) = 1.96

If computing p-value, say test statistic $Z = 2.3$

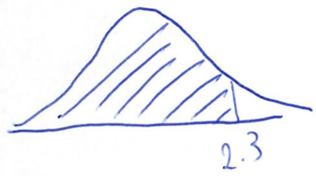


IF ONE SIDED

I need to compute this area

1 - pnorm(2.3)

because pnorm(2.3)



IF TWO SIDED

I need to compute the sum of these two areas

either: 2 * (1 - pnorm(2.3))

or

2 * (~~[illegible]~~ pnorm(-2.3))

qt / pt
qnorm / pnorm

equivalent as they are both symmetric distributions centered at 0

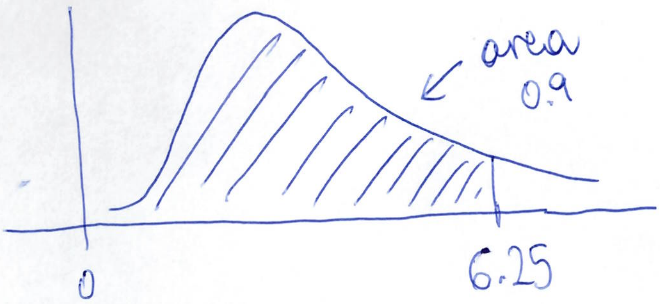
* must define the degrees of freedom for the t distribution.

---

pchisq / qchisq
pf / qf

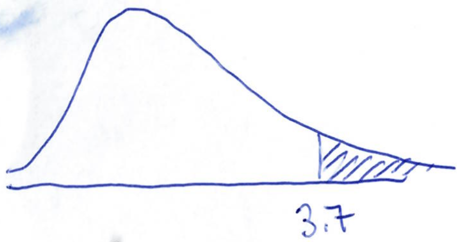
equivalent
→ asymmetric
→ positive only

* must define 2 degrees of freedom parameters for the F distribution



qchisq (0.9 , 3) = 6.25
pchisq (6.25, 3) = 0.9

If computing p-value, say test statistic $W = 3.7$ and 3 degrees of freedom



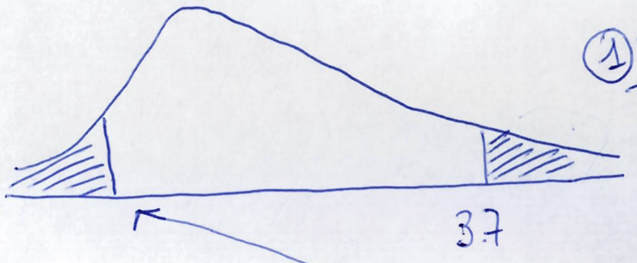
IF ONE SIDED

← I need to compute this area

1 - pchisq (3.7 , 3)

IF TWO SIDED

p.7 : $2 \min\{P(W > W_{obs}), P(W < W_{obs})\}$



1 STEP
TO COMPUTE
AREA

2 STEP
COMPUTE SAME
AREA ON OTHER
SIDE

3 SUM THEM UP

OR JUST MULTIPLY
THIS AREA
BY 2

3.5 test of two proportions

n_1 men

n_2 women

"Was the budget good for the country?"

x_1 men say yes

x_2 women say yes

$\left[\begin{array}{l} n_1 - x_1 \text{ men say no} \\ n_2 - x_2 \text{ women say no} \end{array}\right]$

this can be used to estimate proportions:

$$\hat{p}_1 = \frac{x_1}{n_1} \qquad \hat{p}_2 = \frac{x_2}{n_2}$$

$$H_0 : p_1 = p_2$$

If n_1 and n_2 are large, by the central limit theorem, the distribution is Normal:

$$\hat{p}_1 - \hat{p}_2 \sim N\left(p_1 - p_2, \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}\right)$$

If Ho is true: $p_1 = p_2 = p$

I can estimate $\hat{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$

the test statistic becomes

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim N(0,1)$$

EXAMPLE 3.12

$n_1 = 1{,}000$ $x_1 = 450$

$n_2 = 950$ $x_2 = 390$

$H_0 : p_1 = p_2$ $H_1 : p_1 \neq p_2$

$$\hat{p}_1 = \frac{450}{1000} = 0.45$$

$$\hat{p}_2 = \frac{390}{950} = 0.4105$$

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{390+450}{1000+950} = 0.4308$$

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = 1.759$$

rejection region: qnorm (0.025)

$= 1.96$

3.6 CORRELATION

ρ POPULATION CORRELATION BETWEEN TWO RANDOM VARIABLES

r sample correlation

So,

$$Z' = \frac{1}{2} \log\left(\frac{1+r}{1-r}\right) \approx N\left(\frac{1}{2}\log\left(\frac{1+\rho}{1-\rho}\right), \frac{1}{n-3}\right)$$

only valid for $n > 50$.

$$\frac{\frac{1}{2}\log\left(\frac{1+r}{1-r}\right) - \frac{1}{2}\log\left(\frac{1+\rho}{1-\rho}\right)}{\sqrt{1/n-3}} \sim N(0,1)$$

EXAMPLE 3.14

$H_0 : \rho = 0.5$
$H_1 : \rho \neq 0.5$

$$Z' = \frac{1}{2} \log\left(\frac{1+r}{1-r}\right) = 0.3541$$

$$\frac{1}{2} \log \frac{1+\rho}{1-\rho} = \frac{1}{2} \log\left(\frac{1.5}{0.5}\right) = 0.5493$$

$$\sqrt{1/(n-3)} = \sqrt{1/(84-3)} = 0.1111$$

So the test statistic is

$$\frac{Z' - 0.5493}{0.1111} = \frac{0.3541 - 0.5493}{0.1111} = -1.76$$

TEST TWO CORRELATION COEFFICIENTS

$H_0 : \rho_1 = \rho_2$

So Z'_i will have the same mean in each population. We are testing $Z'_1 - Z'_2$ has mean 0.

Assuming independence of the two samples,

$$Var(Z'_1 - Z'_2) = Var(Z'_1) + Var(Z'_2)$$

$$= \frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}$$

$$= \frac{1}{75-3} + \frac{1}{63-3} = 0.0306$$

the test statistic

$$\frac{(Z'_1 - Z'_2) - 0}{\sqrt{Var(Z'_1 - Z'_2)}} \sim N(0,1)$$

$$\frac{0.8107 - 0.4497}{\sqrt{0.0306}} = 2.064$$