Vengono sottoposti a confronto i consumi delle autovetture Citroen Saxo 1.1 SPI e VW Polo 1.0 alla velocità costante di 120 Km/h. Si ritiene che i consumi dei due tipi di autovetture possa essere descritto da variabili aleatorie con distribuzione normale con la stessa varianza (cioé possiamo assumere $\sigma_C^2 = \sigma_P^2$). La Polo in 20 prove consuma mediamente 6.5 l/100Km, la Saxo in 22 prove consuma mediamente 6.6 l/100Km. Le relative varianze campionarie sono rispettivamente di 0.30 e 0.28.

$$X_c \sim N(\mu_c ; \sigma_c^2) \qquad n_c = 22 \qquad \bar{x}_c = 6{,}6 \qquad s_c^2 = 0{,}28$$
$$X_p \sim N(\mu_p ; \sigma_p^2) \qquad n_p = 20 \qquad \bar{x}_p = 6{,}5 \qquad s_p^2 = 0{,}30$$
$$\sigma_c^2 = \sigma_p^2$$

a) E' ragionevole supporre che le due autovetture abbiano consumi medi differenti al livello di significatività del 5%? $\alpha = 0{,}05$

CASO: TEST SULLA MEDIA DI DUE POPOLAZIONI NORMALI INDIPENDENTI CON VARIANZE INCOGNITE MA UGUALI (OMOSCHEDASTICITÀ)

IPOTESI: $H_0 : \mu_c = \mu_p \qquad H_1 : \mu_c \neq \mu_p$

STATISTICA: $$\hat{t} = \frac{\bar{x}_c - \bar{x}_p}{s\sqrt{\frac{1}{n_c} + \frac{1}{n_p}}} \qquad s^2 = \frac{(n_c-1)s_c^2 + (n_p+1)s_p^2}{n_c + n_p - 2}$$

$$s^2 = \frac{(22-1)\cdot 0{,}28 + (20-1)\cdot 0{,}30}{22+20-2} \Rightarrow s = \sqrt{\frac{21\cdot 0{,}28 + 19\cdot 0{,}30}{40}} \simeq 0{,}5381$$

$$|\hat{t}| = \frac{6{,}6 - 6{,}5}{0{,}5381\sqrt{\frac{1}{22} + \frac{1}{20}}} \simeq 0{,}612$$

REGIONE RIFIUTO: $R = I^c$ CON $I = [-t_\beta(k) ; +t_\beta(k)]$

$$\beta = 1 - \frac{\alpha}{2} = 0{,}975 \qquad k = n_c + n_p - 2 = 40 \qquad t_{0{,}975}(40) \simeq 2{,}021$$

RIFIUTO SE: $|\hat{t}| > t_{0{,}975}(40) \Rightarrow 0{,}612 > 2{,}021$ FALSO

$\Rightarrow$ NON POSSO RIFIUTARE H_0

b) Si calcoli un intervallo di confidenza di livello 95% per la differenza dei consumi medi.

$$x_c - x_p \pm \overset{(*)}{t_\beta(k)}\, s\sqrt{\frac{1}{n_c} + \frac{1}{n_p}} \simeq 0{,}1 \pm 2{,}021 \cdot 0{,}5381 \cdot \sqrt{\frac{1}{22} + \frac{1}{20}} \simeq$$

$$\simeq 0{,}1 \pm 0{,}33586 \Rightarrow IdC : (-0{,}23586 ; +0{,}43586)$$

(*) ATTENZIONE ORA $\alpha = 0{,}95$ MA $\beta = \frac{1+\alpha}{2}$ VALE SEMPRE 0,975

La concentrazione di zinco nell'acqua potabile ne influenza il sapore e può anche risultare nociva. Viene condotta un'analisi in sei punti diversi di un fiume e in ciascun punto viene misurata la concentrazione di zinco in superficie e in profondità. I dati sono mostrati nella seguente tabella: I DATI PROVENGONO DA POP. NORMALI.

	Zona 1	Zona 2	Zona 3	Zona 4	Zona 5	Zona 6
Concentrazione in profondità	0.430	0.266	0.567	0.531	0.707	0.716
Concentrazione in superficie	0.415	0.238	0.390	0.410	0.605	0.609

X_P
X_S

I dati suggeriscono che la vera concentrazione media di zinco in profondità eccede quella in superficie?

$\mu_P > \mu_S$

CASO: TEST SULLA MEDIA DI DUE POPOLAZIONI NORMALI PAIRED CON VARIANZE INCOGNITE

X_P, X_S NORMALI $\Rightarrow X = X_P - X_S \sim N(\mu, \sigma^2)$

i	1	2	3	4	5	6	Σ
X_i	0,015	0,028	0,177	0,121	0,102	0,107	0,550
$\lvert x_i - \bar{x}_n \rvert$	0,077	0,064	0,085	0,029	0,010	0,015	
^2	0,0059	0,0041	0,0073	0,0009	0,0001	0,0002	0,01842

$$\bar{x}_n = \frac{1}{6}\sum_1^6 x_i = \frac{0,550}{6} = 0,091\bar{6}$$

$$S^2 = \frac{1}{5}\sum_1^6 (x_i - \bar{x}_n)^2 = \frac{0,01842}{5} \simeq 0,003683 \Rightarrow S \simeq 0,060688$$

IPOTESI $\quad H_0: \mu_P - \mu_S \leq 0 \quad H_1: \mu_P - \mu_S > 0$

STATISTICA $\quad \hat{t} = \frac{\bar{x}_n}{S}\sqrt{n} = \frac{0,0916}{0,060688}\sqrt{6} \simeq 3,7$

REGIONE RIFIUTO $\quad R = [t_\beta(n-1); +\infty) \quad \beta = 1-\alpha = 0,95$

$t_{0,95}(5) = 2,01505$

RIFIUTO SE $\quad \hat{t} > t_\beta(n-1) \Rightarrow 3,7 > 2,01505$ VERO

$\Rightarrow$ RIFIUTO H_0

CON $\alpha = 0,01 \Rightarrow t_\beta(n-1) = t_{0,99}(5) \simeq 3,365$

CONTINUO A RIFIUTARE H_0

I dati x_i esposti nella tabella seguente mostrano $n=10$ misurazioni della concentrazione di iodio nella stessa soluzione. Al fine di poter giudicare la precisione delle misurazioni, nell'ipotesi che i dati provengano da una popolazione gaussiana, si costruisca un intervallo fiduciario al 95% per σ^2.

Prova	Concentrazione	Prova	Concentrazione
1	5.507	6	5.527
2	5.506	7	5.504
3	5.500	8	5.490
4	5.497	9	5.500
5	5.506	10	5.497

(Mendenhall e Sincich, 1989, pag. 308)

i	x	x-x_n	(x-x_n)^2
1	5,507	0,004	0,000013
2	5,506	0,003	0,000007
3	5,500	-0,003	0,000012
4	5,497	-0,006	0,000041
5	5,506	0,003	0,000007
6	5,527	0,024	0,000557
7	5,504	0,001	0,000000
8	5,490	-0,013	0,000180
9	5,500	-0,003	0,000012
10	5,497	-0,006	0,000041
sum	55,034		**0,000868**
mean	5,503		

L'IdC PER σ^2 CON μ INCOGNITA È DATO DA:

$$I \equiv \left(\frac{(n-1)S_n^2}{\chi^2_{1-\beta}(n-1)} ; \frac{(n-1)S_n^2}{\chi^2_{\beta}(n-1)} \right) \quad \text{DOVE}$$

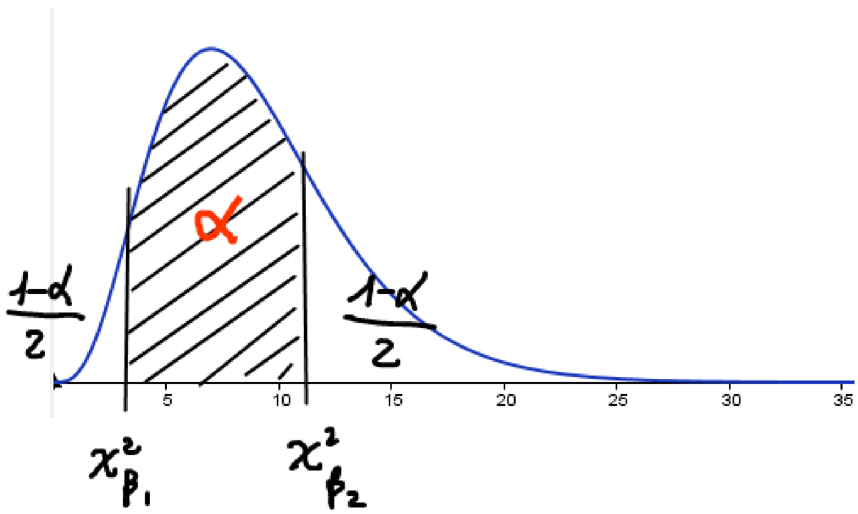
$$n=10 \; ; \; (n-1)S_n^2 = \sum_{i=1}^{n} (x_i - \bar{x}_n)^2 = 0{,}000868$$

$$\beta = \frac{1-\alpha}{2} = \frac{1-0{,}95}{2} = 0{,}025 \qquad 1-\beta = 0{,}975$$

$$\chi^2_{0{,}025}(9) = 2{,}70 \quad ; \quad \chi^2_{0{,}975}(9) = 19{,}02$$

$$I \equiv \left(\frac{0{,}000868}{19{,}02} ; \frac{0{,}000868}{2{,}70} \right) = (0{,}0000457 ; 0{,}0003216)$$

$$A = 0{,}0002759$$



$$\beta_1 = \frac{1-\alpha}{2} = \beta \qquad \beta_2 = 1 - \frac{1-\alpha}{2} = \frac{1+\alpha}{2} = 1-\beta$$

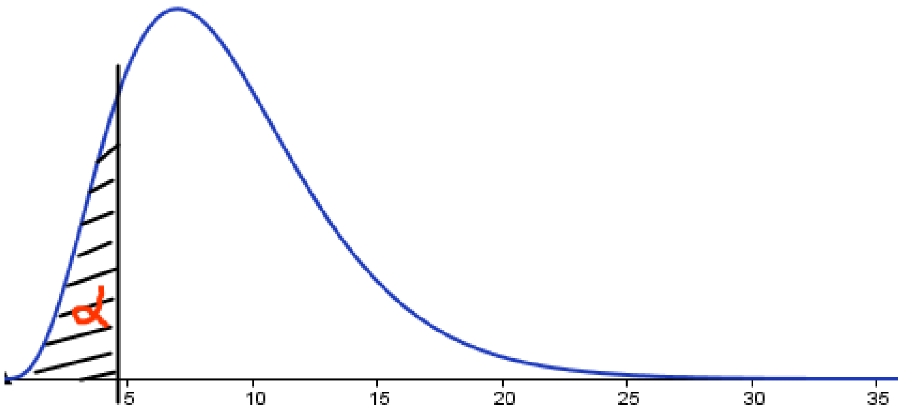
Sottoporre a test con significatività pari 0.05 l'ipotesi $\sigma^2 < 0.001$

CASO: TEST PER σ^2 CON μ INCOGNITA POP. NORMALE

IPOTESI: $H_0: \sigma^2 \geq 0{,}001 \quad H_1: \sigma^2 < 0{,}001$

REGIONE $R = \left[0;\ \chi^2_\beta(n-1)\right]$ $\quad n-1 = 9 \quad \beta = \alpha = 0{,}05$

$$\chi^2_{0{,}05}(9) = 3{,}325$$



$\chi^2_\beta \qquad \beta = \alpha$

STATISTICA: $W_n = \dfrac{(n-1)S_n^2}{\sigma_0^2} = \dfrac{0{,}000868}{0{,}001} = 0{,}868$

RIFIUTO SE: $W_n < \chi^2_\beta(n-1) \qquad 0{,}868 < 3{,}325$ VERO

$\Rightarrow$ RIFIUTO H_0 $\Rightarrow$ LA VARIANZA NON ECCEDE 0,001

Calcolare il p-value

$\bar{\alpha}: \chi^2_{\bar{\alpha}}(n-1) = W_n$ CIOÈ $\chi^2_{\bar{\alpha}}(9) = 0{,}868$

DA CUI CON L'USO DEL CALCOLATORE: $\bar{\alpha} \simeq 0{,}000314$

```
> pchisq(0.868,9)
[1] 0.0003137562
```

Come cambiano i risultati precedenti nel caso la media sia nota e pari a $\mu = 5.5$?

IdC BILATERO PER σ^2 CON μ NOTA

$$I \equiv \left(\frac{n T_n^2}{\chi^2_{1-\beta}(n)} ; \frac{n T_n^2}{\chi^2_{\beta}(n)} \right) \qquad n = 10 \quad \beta = 0{,}025$$

$$n T_n^2 = \sum_{i=1}^{n} (x_i - \mu)^2 \simeq 0{,}000984$$

$$\chi^2_{0,025}(10) = 3{,}247 \qquad \chi^2_{0,975}(10) = 20{,}483$$

$$I \equiv \left(\frac{0{,}000984}{20{,}483} ; \frac{0{,}000984}{3{,}247} \right) \qquad I \equiv (0{,}000048 ; 0{,}0003028)$$

$A = 0{,}0002547$ L'AMPIEZZA SI RIDUCE

Sottoporre a test con significatività pari 0.05 l'ipotesi $\sigma^2 < 0.001$

CASO : TEST PER σ^2 CON μ NOTA

IPOTESI : $H_0 : \sigma^2 \geq 0{,}001$ $\quad H_1 : \sigma^2 < 0{,}001$

REGIONE $R = [0 ; \chi^2_{\beta}(n)]$ $\qquad n = 10 ; \beta = \alpha = 0{,}05$

$$\chi^2_{0,05}(10) = 3{,}94$$

STATISTICA : $$W_n = \frac{n T_n^2}{\sigma_0^2} = \frac{0{,}000984}{0{,}001} = 0.984$$

RIFIUTO SE : $W_n < \chi^2_{\beta}(n)$ $\qquad 0{,}984 < 3{,}94$ VERO

$\Rightarrow$ RIFIUTO H_0 $\Rightarrow$ LA VARIANZA NON ECCEDE 0,001

Calcolare il p-value

$\bar{\alpha} : \chi^2_{\bar{\alpha}}(n) = W_n$ CIOÈ $\chi^2_{\bar{\alpha}}(10) = 0{,}984$

DA CUI CON L'USO DEL CALCOLATORE : $\bar{\alpha} \simeq 0{,}00016$

```
> pchisq(0.984,10)
[1] 0.000159829
```