ESERCIZIO 7.1

Un nuovo modello di termostato per frigorifero dovrebbe assicurare, stando alle specifiche tecniche, una minore variabilità nella temperatura del frigo rispetto ai modelli della concorrenza. In particolare il termostato promette di essere in grado di assicurare una variabilità inferiore a 0.1 (σ), che rappresenta l'attuale standard di mercato per tali apparecchiature. Tale dispositivo viene montanto su 10 frigoriferi sui quali si rileva una deviazione standard pari a 0.04. I dati campionari sono tali da affermare che la deviazione standard reale di tale tipo di termostati è minore di 0.1?

a) Si costruisca un test di ipotesi utilizzando un livello di significatività del 5%.
b) Si calcoli inoltre il livello di significatività osservato (valore p) del test.

SVOLGIMENTO

a)

1) Ipotesi nulla

Ho: $\sigma^2 = 0.01$

2) Ipotesi alternativa

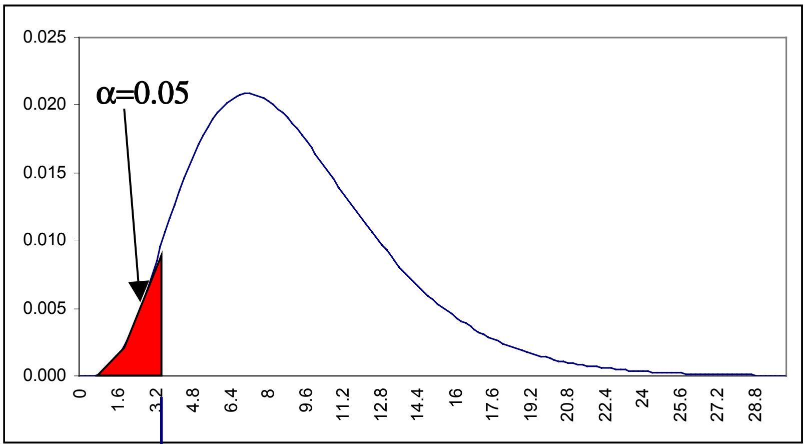
Ha: $\sigma^2 < 0.01$

3) Statistica test

Statistica test → $\dfrac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$

4) Regola di decisione

α=0.05 → $\chi^2_{n-1,1-\alpha} = \chi^2_{10-1,0.95} = 3.3251$



⇒ Rifiuto Ho se la statistica test è minore di 3.3251

3.32
Regione di rifiuto

5) Assunzioni (ipotesi mantenute)

La variabile temperatura del frigo segue una distribuzione normale

6) Esperimento sul campione

$s^2 = 0.4$ n=10

Statistica test → $\dfrac{(n-1)s^2}{\sigma^2} = \dfrac{(10-1)0.04^2}{0.1^2} = 1.44$

7) Conclusione

La statistica test, per il campione considerato, produce un valore nella regione di accettazione: i dati empirici permettono di accettare l'ipotesi che l'adozione del nuovo termostato comporta una varianza minore di 0.01.

b)

Il livello di significatività osservato del test, noto anche come p-value o valore p, è la probabilità (sotto l'ipotesi che H_0 sia vera) di osservare un valore che sia almeno tanto contraddittorio rispetto all'ipotesi nulla e tale da avvalorare l'ipotesi alternativa quanto quello calcolato sui dati campionari a disposizione. E' possibile cercare il valore di χ^2=1.44 sulle tavole della distribuzione chi-quadro facendo riferimento alla riga relativa a 9 gradi di libertà:

P($X^2 > \chi^2$)

	0.995	**0.990**	**0.975**	**0.950**	**0.900**	**0.100**	**0.050**	**0.025**	**0.010**	**0.005**
1	0.00004	0.00016	0.00098	0.00393	0.01579	2.70554	3.84146	5.02390	6.63489	7.87940
2	0.01002	0.02010	0.05064	0.10259	0.21072	4.60518	5.99148	7.37778	9.21035	10.59653
3	0.07172	0.11483	0.21579	0.35185	0.58438	6.25139	7.81472	9.34840	11.34488	12.83807
4	0.20698	0.29711	0.48442	0.71072	1.06362	7.77943	9.48773	11.14326	13.27670	14.86017
5	0.41175	0.55430	0.83121	1.14548	1.61031	9.23635	11.07048	12.83249	15.08632	16.74965
6	0.67573	0.87208	1.23734	1.635[illegible]	2.20413	10.64464	12.59158	14.44935	16.81187	18.54751
7	0.98925	1.23[illegible]	[illegible]	[illegible]	[illegible]	[illegible]03	14.06713	16.01277	18.47532	20.27774
8	1.34440	1.64[illegible]	[illegible]	[illegible]	[illegible]	[illegible]56	15.50731	17.53454	20.09016	21.95486
9	1.73491	2.08789	2.70039	3.32512	4.16816	14.68366	16.91896	19.02278	21.66605	23.58927
10	2.15585	2.55820	3.24696	3.94030	4.86518	15.98717	18.30703	20.48320	23.20929	25.18805

Area che lascia a destra 0.995

Da cui si vede come il valore che più si avvicina a 1.44 è il valore 1.73, che lascia a destra ben 0.995: il test risulta quindi significativo, confermando quanto ottenuto con la procedura standard seguita al passo a.

ESERCIZIO 7.2

Usando i dati dell'esercizio precedente costruire un intervallo di confidenza al 95% per la varianza della popolazione.

SVOLGIMENTO

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$$

$$P\left(\chi^2_{\left(1-\alpha/2\right),n-1} \le \frac{(n-1)s^2}{\sigma^2} \le \chi^2_{\alpha/2,n-1}\right) = 100\times(1-\alpha)$$

$$P\left(\frac{\chi^2_{\left(1-\alpha/2\right),n-1}}{(n-1)s^2} \le \frac{1}{\sigma^2} \le \frac{\chi^2_{\alpha/2,n-1}}{(n-1)s^2}\right) = 100\times(1-\alpha)$$

$$P\left(\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{\left(1-\alpha/2\right),n-1}}\right) = 100\times(1-\alpha)$$

Sostituendo i valori dell'esercizio si ha:

1)

$$\alpha = 0.05 \Rightarrow 100\times(1-\alpha) = 0.95$$

$$P\left(\frac{(10-1)0.04^2}{\chi^2_{\alpha/2,n-1}} \le \sigma^2 \le \frac{(10-1)0.04^2}{\chi^2_{\left(1-\alpha/2\right),n-1}}\right) = 100\times(1-\alpha)$$

$$P\left(\frac{(10-1)0.04^2}{\chi^2_{0.025,9}} \le \sigma^2 \le \frac{(10-1)0.04^2}{\chi^2_{0.975,9}}\right) = 95\%$$

Dalle tavole della distribuzione chi-quadro è possibile trovare i due valori di interesse:

$P(X^2 > \chi^2)$

	0,995	0,990	0,975	0,950	0,900	0,100	0,050	0,025	0,010	0,005
1	0,00004	0,00016	0,00098	0,00393	0,01579	2,70554	3,84146	5,02390	6,63489	7,87940
2	0,01002	0,02010	0,05064	0,10259	0,21072	4,60518	5,99148	7,37778	9,21035	10,59653
3	0,07172	0,11483	0,21579	0,35185	0,58438	6,25139	7,81472	9,34840	11,34488	12,83807
4	0,20698	0,29711	0,48442	0,71072	1,06362	7,77943	9,48773	11,14326	13,27670	14,86017
5	0,41175	0,55430	0,83121	1,14548	1,61031	9,23635	11,07048	12,83249	15,08632	16,74965
6	0,67573	0,87208	1,23734	1,63538	2,20413	10,64464	12,59158	14,44935	16,81187	18,54751
7	0,98925	1,23903	1,68986	2,16735	2,83311	12,01703	14,06713	16,01277	18,47532	20,27774
8	1,34440	1,64651	2,17972	2,73263	3,48954	13,36156	15,50731	17,53454	20,09016	21,95486
9	1,73491	2,08790	2,70039	3,32512	4,16816	14,68366	16,91898	19,02278	21,66605	23,58927

$\chi^2_{0.975,9} \rightarrow 2.70039$ $\qquad$ $\chi^2_{0.025,9} \rightarrow 19.02278$

Sostituendo tali valori nella formula si ha:

$$P\left(\frac{(10-1)0.04^2}{19.02278} \le \sigma^2 \le \frac{(10-1)0.04^2}{2.70039}\right) = 95\%$$

$$P\left(0.000757 \le \sigma^2 \le 0.005333\right) = 95\%$$

ESERCIZIO 7.3

Un economista del Ministero degli Esteri desidera verificare se gli accordi di negoziazione tra Italia e Giappone siano rispettati. In particolare egli sospetta che i produttori giapponesi fissino un prezzo più basso per i prodotti venduti sul mercato italiano rispetto a quello usato sul mercato interno, ostacolando al contempo le importazioni di prodotti italiani con forti ostacoli di tipo burocratico. Si interessa in particolare al mercato dell'auto e vuole testare l'ipotesi che prezzi più alti siano applicati in Giappone rispetto all'Italia per le autovetture di produzione giapponese. Esamina a tal fine due campioni relativi a pratiche di acquisto di tali autovetture nello stesso periodo di tempo (50 per il mercato italiano e 30 per il mercato giapponese). Convertendo i prezzi di vendita in Giappone usando il cambio corrente Yen/Euro, ottiene i risultati elencati nella seguente tabella:

	ITALIA	GIAPPONE
Ampiezza campione	50	30
Media campionaria	€ 16545	€ 17243

Siano inoltre noti i seguenti valori per le rispettive popolazioni di riferimento:

	ITALIA	GIAPPONE
Deviazione standard	€ 1989	€ 1843

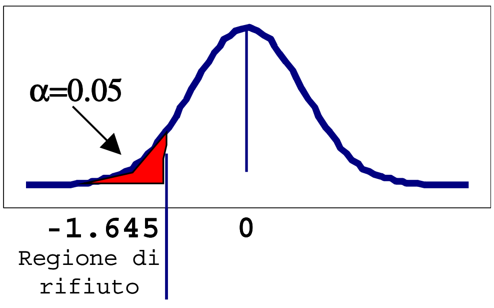
a) Costruire un test di ipotesi usando un livello α=0.5
b) Si calcoli inoltre il livello di significatività osservato del test (p-value)

SVOLGIMENTO

a)
1) Ipotesi nulla
Ho: $(\mu_1 - \mu_2) = 0$
2) Ipotesi alternativa
Ha: $(\mu_1 - \mu_2) < 0$
3) Statistica test

Statistica test → $$\frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{(\bar{x}_1 - \bar{x}_2)}} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim Z$$

4) Regola di decisione
Utilizzando le tavole della Z si ha infatti:
α=0.05 → $Z_\alpha = -1.645$



⇒ Rifiuto Ho se la statistica test è minore di −1.645

5) Assunzioni (ipotesi mantenute)
I due campioni sono selezionati in maniera indipendente dalle due popolazioni. Le rispettive ampiezze campionarie n_1 e n_2 sono sufficientemente grandi affinché sia $\bar{x}_1$ che $\bar{x}_2$ siano distribuite approssimativamente come normali.

6) Esperimento sul campione

$\bar{x}_1 = 16545$	$\sigma_1 = 1989$	n_1=50
$\bar{x}_2 = 17243$	$\sigma_2 = 1843$	n_2=30

Statistica test → $\dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} = \dfrac{(16545 - 17243)}{\sqrt{\dfrac{1989^2}{50} + \dfrac{1843^2}{30}}} = -1.59$

7) Conclusione
La statistica test, per il campione considerato, produce un valore nella regione di accettazione: i dati empirici non permettono di accettare l'ipotesi che il prezzo praticato in Giappone per le autovetture in questione sia più alto di quello praticato in Italia.

b)
Il livello di significatività osservato del test, noto anche come p-value o valore p, è la probabilità (sotto l'ipotesi che H_0 sia vera) di osservare un valore che sia almeno tanto contraddittorio rispetto all'ipotesi nulla e tale da avvalorare l'ipotesi alternativa quanto quello calcolato sui dati campionari a disposizione.
E' possibile cercare il valore di Z=-1.56 dalle tavole della normale standardizzata:

P(0 < Z < z)

	0.03	**0.04**	**0.05**	**0.06**
1.2	0.3907	0.3925	0.3944	0.3962
1.3	0.4082	0.4099	0.4115	0.4131
1.4	0.4236	0.4251	0.4265	0.4279
1.5	0.4370	0.4382	0.4394	0.4406

Area tra la media e il valore z

Da cui:
$P(X < 4.5) = P(Z < -1.56) = P(Z > 1.56) = 0.5 - 0.4406 = 0.06$

Fissando α=0.05 (probabilità errore di prima specie) il livello di significatività osservato non è tale da rifiutare H_0 (come visto anche con la procedura standard seguita al passo a).

ESERCIZIO 7.4

Si ripeta l'esercizio 7.3 nell'ipotesi che non vi siano informazioni sulle varianze delle due popolazioni. Si utilizzino a tal fine i seguenti valori campionari:

	ITALIA	GIAPPONE
Deviazione standard	€ 1989	€ 1843

SVOLGIMENTO

a)

1) Ipotesi nulla

Ho: $(\mu_1 - \mu_2) = 0$

2) Ipotesi alternativa

Ha: $(\mu_1 - \mu_2) < 0$

3) Statistica test

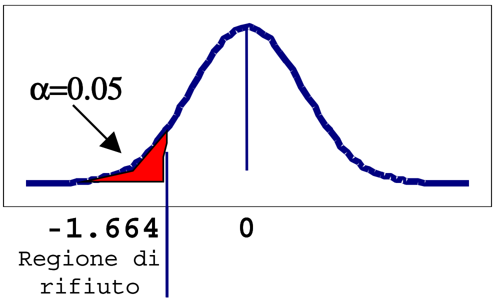
Statistica test → $\dfrac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{\sigma_{(\overline{x}_1 - \overline{x}_2)}} = \dfrac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim Z$

Non conoscendo la varianza delle due popolazioni è necessario utilizzare la distribuzione t di Student per la statistica test sotto riportata:

$$\frac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_{x_1}^2 (n_1 - 1) + s_{x_2}^2 (n_2 - 1)}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} \sim t_{n_1 + n_2 - 2}$$

4) Regola di decisione

$\alpha = 0.05$ → $t_{n_1 + n_2 - 2, \alpha} = -1.664$



⇒ Rifiuto Ho se la statistica test è minore di −1.664

5) Assunzioni (ipotesi mantenute)

I due campioni sono selezionati in maniera indipendente dalle due popolazioni. Le rispettive ampiezze campionarie n_1 e n_2 sono sufficientemente grandi affinché sia $\overline{x}_1$ che $\overline{x}_2$ siano distribuite approssimativamente come normali e s_1^2 e s_2^2 forniscano, rispettivamente, una buona stima di σ_1^2 e σ_2^2.

6) Esperimento sul campione

$\overline{x}_1 = 16545$	$s_1 = 1989$	$n_1 = 50$
$\overline{x}_2 = 17243$	$s_2 = 1843$	$n_2 = 30$

Statistica test → $\dfrac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_{x_1}^2 (n_1 - 1) + s_{x_2}^2 (n_2 - 1)}} \sqrt{\dfrac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} \sim t_{n_1 + n_2 - 2}$

$$\frac{(16545-17243)}{\sqrt{1989^2(50-1)+1843^2(30-1)}}\sqrt{\frac{50\times 30(50+30-2)}{50+30}} \sim t_{50+30-2}$$

$$\frac{-698}{17098.326}\times 38.24=-1.56$$

7) Conclusione

La statistica test, per il campione considerato, produce un valore nella regione di accettazione: i dati empirici non permettono di accettare l'ipotesi che il prezzo praticato in Giappone per le autovetture in questione sia più alto di quello praticato in Italia.

b)

Il livello di significatività osservato del test, noto anche come p-value o valore p, è la probabilità (sotto l'ipotesi che H_0 sia vera) di osservare un valore che sia almeno tanto contraddittorio rispetto all'ipotesi nulla e tale da avvalorare l'ipotesi alternativa quanto quello calcolato sui dati campionari a disposizione.

Si può ottenere un valore approssimato per la t_{25} usando le tavole:

	0,100	**0,050**	**0,025**	**0,020**	**0,010**	**0,005**	**0,0025**	**0,001**	**0,0005**
1	3,078	6,314	12,706	15,894	31,821	63,656	127,321	318,289	636,578
2	1,886	2,920	4,303	4,849	6,965	9,925	14,089	22,328	31,600
3	1,638	2,353	3,182	3,482	4,541	5,841	7,453	10,214	12,924
...	...	...	...	...	...	...	...	...	...
80	1,292	1,664	1,990	2,088	2,374	2,639	2,887	3,195	3,416
100	1,290	1,660	1,984	2,081	2,364	2,626	2,871	3,174	3,390
120	1,289	1,658	1,980	2,076	2,358	2,617	2,860	3,160	3,373
INF	1,282	1,645	1,960	2,054	2,326	2,576	2,807	3,090	3,290

Dalle tavole si può notare come il valore cercato corrisponda a un p-value compreso tra 0.100 e 0.050: il test risulta quindi poco significativo rispetto al livello α fissato, confermando quanto ottenuto con la procedura standard seguita al passo a.

ESERCIZIO 7.5

Un produttore di carta vuole confrontare la variazione nei livelli di produzione giornalieri in due cartiere di sua proprietà. Seleziona casualmente campioni casuali indipendenti relativi a differenti giorni di produzione registrando per ciascuno il numero di pezzi prodotti. I dati sintetici di questa rilevazione sono riportati nella seguente tabella:

	Cartiera 1	Cartiera 2
Ampiezza campione	13	18
Media campionaria	26.3	19.7
Deviazione standard	8.2	4.7

a) Costruire un test di ipotesi usando un livello α=0.10
b) Si calcoli inoltre il livello di significatività osservato del test (p-value)

SVOLGIMENTO

a)

1) Ipotesi nulla

Ho: $\dfrac{\sigma_1^2}{\sigma_2^2}=1 \qquad \left(\sigma_1^2=\sigma_2^2\right)$

2) Ipotesi alternativa

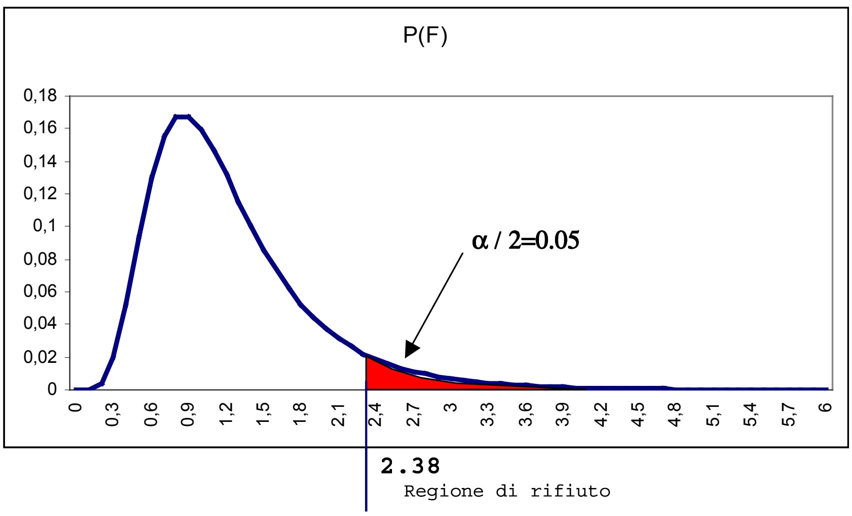
Ha: $\dfrac{\sigma_1^2}{\sigma_2^2}\neq 1 \qquad \left(\sigma_1^2\neq\sigma_2^2\right)$

3) Statistica test

Statistica test → $\dfrac{s_1^2}{s_2^2}\sim F_{n_1-1,n_2-1}$

4) Regola di decisione

α=0.10 → $F_{(n_1-1,n_2-1),\alpha/2}=F_{(12,17),0.05}=2.38$



⇒ Rifiuto Ho se la statistica test è maggiore di 2.38

5) Assunzioni (ipotesi mantenute)

- I due campioni sono selezionati in maniera indipendente dalle due popolazioni.
- Le due variabili, livello di produzione giornaliera in ciascuna delle due cartiere sono distribuite normalmente.

6) Esperimento sul campione

$\bar{x}_1 = 26.3$ $\quad s_1 = 8.2$ $\quad n_1=13$

$\bar{x}_2 = 19.7$ $\quad s_2 = 4.7$ $\quad n_2=18$

Statistica test → $\dfrac{\text{varianza campionaria maggiore}}{\text{varianza campionaria minore}} = \dfrac{s_1^2}{s_2^2} = \dfrac{8.2^2}{4.7^2} = 3.04$

7) Conclusione

La statistica test, per il campione considerato, produce un valore nella regione di rifiuto: i dati empirici permettono di accettare l'ipotesi che vi sia differenza nella variabilità della produzione giornaliera delle due cartiere.

b)

Il livello di significatività osservato del test, noto anche come p-value o valore p, è la probabilità (sotto l'ipotesi che H_0 sia vera) di osservare un valore che sia almeno tanto contraddittorio rispetto all'ipotesi nulla e tale da avvalorare l'ipotesi alternativa quanto quello calcolato sui dati campionari a disposizione. E' possibile cercare il valore di F=3.04 dalle tavole della distribuzione F di Fisher. Considerando una $F_{12,17}$ il valore per α=0.025 è 2.82 mentre il valore per α=0.01 è 3.46. Il valore del p-value può essere quindi approssimato come segue:

$0.01 \times 2 <$ p-value $< 0.025 \times 2$ (è necessario raddoppiare α perchè è un test bilaterale)

$0.02 <$ p-value < 0.05

Confermando quindi quanto ottenuto con la procedura standard seguita al passo a.

ESERCIZIO 7.6

Si ripeta l'esercizio 7.5 ipotizzando che le medie campionarie siano uguali alle medie note delle rispettive popolazioni.

SVOLGIMENTO

a)

1) Ipotesi nulla

Ho: $\frac{\sigma_1^2}{\sigma_2^2}=1 \qquad \left(\sigma_1^2=\sigma_2^2\right)$

2) Ipotesi alternativa

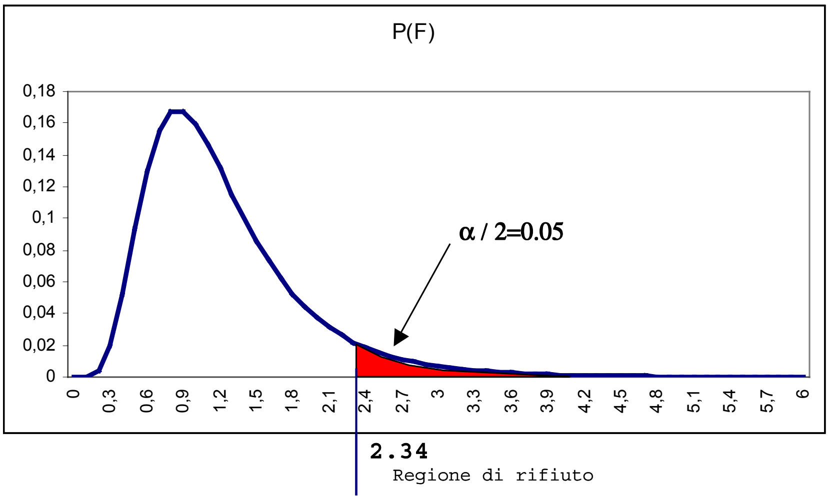
Ha: $\frac{\sigma_1^2}{\sigma_2^2}\neq 1 \qquad \left(\sigma_1^2\neq\sigma_2^2\right)$

3) Statistica test

Statistica test → $\dfrac{\frac{1}{n_1}\sum_{i=1}^{n}\left(x_{1,i}-\mu\right)^2}{\frac{1}{n_2}\sum_{i=1}^{n}\left(x_{2,i}-\mu\right)^2} \sim F_{n_1,n_2}$

4) Regola di decisione

α=0.10 → $F_{(n_1,n_2),\alpha/2}=F_{(13,18),0.05}=2.34$



⇒ Rifiuto Ho se la statistica test è maggiore di 2.34

5) Assunzioni (ipotesi mantenute)

- I due campioni sono selezionati in maniera indipendente dalle due popolazioni.
- Le due variabili, livello di produzione giornaliera in ciascuna delle due cartiere sono distribuite normalmente.

6) Esperimento sul campione

$\mu_1=26.3 \qquad \hat{\sigma}_1=\frac{1}{n_1}\sum_{i=1}^{n}\left(x_{1,i}-\mu_1\right)^2=8.2 \qquad n_1=13$

$\mu_2=19.7 \qquad \hat{\sigma}_2=\frac{1}{n_2}\sum_{i=1}^{n}\left(x_{2,i}-\mu_2\right)^2=4.7 \qquad n_2=18$

Statistica test → $\frac{\text{varianza campionaria maggiore}}{\text{varianza campionaria minore}} = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} = \frac{8.2^2}{4.7^2} = 3.04$

7) Conclusione

La statistica test, per il campione considerato, produce un valore nella regione di rifiuto: i dati empirici permettono di accettare l'ipotesi che vi sia differenza nella variabilità della produzione giornaliera delle due cartiere.

b)

Il livello di significatività osservato del test, noto anche come p-value o valore p, è la probabilità (sotto l'ipotesi che H_0 sia vera) di osservare un valore che sia almeno tanto contraddittorio rispetto all'ipotesi nulla e tale da avvalorare l'ipotesi alternativa quanto quello calcolato sui dati campionari a disposizione.

E' possibile cercare il valore di F=3.04 dalle tavole della distribuzione F di Fisher. Considerando una $F_{13,18}$ il valore per α=0.025 è 2.77 mentre il valore per α=0.01 è 3.37. Il valore del p-value può essere quindi approssimato come segue:

$0.01 < \text{p-value} < 0.025$

Confermando quindi quanto ottenuto con la procedura standard seguita al passo a.