# Esercitazione #5 di Statistica

# Test ed Intervalli di Confidenza (per una popolazione)

2 Dicembre 2002

## 1 Esercizi

### 1.1 Test su media (con varianza nota)

**Esercizio n. 1**

Il calore (in calorie per grammo) emesso da un composto di cemento è (approssimatamente) normalmente distribuito di deviazione standard nota $\sigma = 2$. Si vuole testare

$$H_0 : \mu = 100$$

contro

$$H_1 : \mu \neq 100$$

con un campione di dimensione $n = 9$.

1. Se la regione di accettazione fosse data da $98.5 \leq \bar{x} \leq 101.5$, quale sarebbe l'errore di prima specie $\alpha$?

2. Determinare l'errore di seconda specie $\beta$ e la potenza del test quando la vera media del calore è pari a 103.

3. Determinare l'errore di seconda specie $\beta$ quando la vera media del calore è pari a 105. Tra il $\beta$ appena trovato e quello trovato nel punto 2., quale dei due è più piccolo e perché?

**Esercizio n. 2**

Un'azienda produce anelli per i pistoni delle automobili. E' noto che il diametro di tali anelli è approssimatamente normalmente distribuito ed ha deviazione standard $\sigma = 0.001$ mm. Da un campione di $n = 15$ anelli, inoltre, si ricava una media campionaria $\bar{x} = 74.036$ mm.

Si vuole testare l'ipotesi che la media del diametro degli anelli sia uguale a 74.035 mm. ad un livello di significatività $\alpha = 5\%$.

1. Costruire la regione di accettazione dell'ipotesi enunciata sopra.
2. Quale conclusione possiamo trarre con i dati a noi disponibili?

**Esercizio 3**

Supponiamo che la vita (in ore) di una lampadina da 75 watt sia approssimatamente normalmente distribuita ed abbia una deviazione standard pari a $\sigma = 25$ ore. Un campione di 20 lampadine ha una media (campionaria) di vita $\bar{x} = 1014$ ore.

1. E' ragionevole supporre che la media di vita delle lampadine superi 1000 ore? Usare un livello di significatività $\alpha = 5\%$.
2. Calcolare il $P - value$ del test precedente.

**Esercizio 4**

Il tasso di un processo chimico è studiato. Di tale tasso sono note la deviazione standard $\sigma = 8$ e la media campionaria $\bar{x} = 88.48\%$, ottenute da un campione di dimensione $n = 100$.

1. Determinare il $P - value$ del test in cui ci si domanda se sia ragionevole che la media del tasso non sia 90%.
2. Quali conclusioni possiamo trarre ai livelli di significatività del 5% e del 6%?

## 1.2 Intervalli di confidenza per la media (con varianza nota)

**Esercizio 5**

Sappiamo che un intervallo di confidenza al 95% per la media di una popolazione normalmente distribuita è dato da $(11.4104; 13.8896)$. Se la varianza è pari a 6:

1. quale era la media campionaria della popolazione ?

2. qual è la dimensione del campione dal quale abbiamo determinato l'intervallo di confidenza in questione?

**Esercizio 6**

La lunghezza (in cm.) delle mine per matita di una certa marca è distribuita come una variabile aleatoria gaussiana. Supponiamo che la precisione dello strumento che produce le mine sia nota e che pertanto la deviazione standard della lunghezza delle mine sia pari a $\sigma = 0.1$.

Misurando dieci mine, otteniamo i seguenti valori:

$$
\begin{array}{lllll}
12.21; & 12.33; & 12.84; & 12.97; & 13.22; \\
12.93; & 13.07; & 13.52; & 13.23; & 13.01
\end{array}
$$

1. Determinare un intervallo di confidenza al 95% per la lunghezza media delle mine.
2. Mantenendo lo stesso livello di confidenza, quante mine occorrerebbe misurare per avere un intervallo di ampiezza inferiore o al più uguale a $10^{-2}$ cm.?

## 1.3 Test su media (con varianza non nota)

**Esercizio 7**

L'etichetta delle bottiglie di champagne Veuve Coquelicot dichiara un contenuto di 730 ml. Un'associazione di consumatori decide di controllare questa affermazione e su 81 bottiglie esaminate riscontra una media campionaria $\bar{x} = 726$ ml. ed una varianza campionaria $s^2 = 625$.

Supponendo che la quantità di champagne contenuta in ogni bottiglia si possa modellizzare con una v.a. normale, si può concludere (al livello di significatività $\alpha = 5\%$) che in media le bottiglie contengono meno di quanto dichiarato?

**Esercizio 8**

L'altezza media delle reclute alla visita di leva nel 1970 era di 169 cm.

121 reclute vengono scelte a caso nel 1980 e da queste vengono trovate una media campionaria $\bar{x} = 171$ cm. e una varianza campionaria $s^2 = 85$. Si può affermare (al livello di significatività $\alpha = 5\%$) che l'altezza media delle reclute è rimasta invariata?

**Esercizio 9**

Nella produzione di semiconduttori non è possibile controllare esattamente la resistenza degli elementi prodotti. Supponiamo che vengano misurati i valori della resistenza per $n = 81$ semiconduttori, ottenendo una media campionaria $\bar{x} = 1.2$ ed una varianza campionaria $s^2 = 0.4$.

1. Determinare l'intervallo bilaterale di confidenza al 95% per la media della resistenza dei semiconduttori prodotti.

2. Al livello di significatività $\alpha = 5\%$, è possibile accettare l'ipotesi nulla

$$H_0 : \mu = 1.3$$

contro

$$H_1 : \mu \neq 1.3 \text{ ?}$$

**Esercizio 10**

Vengono effettuate 20 misurazioni della concentrazione di un certo enzima nel sangue di diversi individui e si osservano una media campionaria $\bar{x} = 1.23$ ed una varianza campionaria $s^2 = 0.4$.

1. Supponendo che i valori di questa concentrazione seguano una distribuzione normale (con entrambi i parametri ignoti!), qual è un intervallo di fiducia al livello 95% per la media della concentrazione?

2. Quale sarebbe l'intervallo di cui sopra se la concentrazione a cui si è interessati fosse distribuita come una normale di varianza nota $\sigma^2 = 0.4$?

3. Quale tra i due intervalli trovati in 1. ed in 2. è più ampio? Il risultato ottenuto sembra ragionevole?

## 1.4 Test su varianza

**Esercizio 11**

La quantità di zucchero contenuta in un succo di pesca è normalmente distribuita.

Si vuole testare l'ipotesi

$$H_0 : \sigma^2 = 18 \; mg.^2$$

contro l'ipotesi alternativa

$$H_1 : \sigma^2 \neq 18 \; mg.^2$$

1. Se da un campione di $n = 10$ succhi di pesca otteniamo una varianza campionaria $s^2 = 23.04 \; mg.^2$, possiamo accettare oppure no l'ipotesi nulla ad un livello $\alpha = 5\%$?

2. Determinare l'intervallo di confidenza bilaterale al 98% per $\sigma$.

3. Con l'aiuto del punto 2., che cosa possiamo concludere circa l'ipotesi nulla ad un livello $\alpha = 2\%$?

**Esercizio 12**

La margarina di una certa marca viene analizzata con lo scopo di determinare il livello di grassi insaturi (in percentuale) in essa contenuti.

Un campione di 6 confezioni fornisce i seguenti risultati:

$$16.6; \quad 17.1; \quad 17.4; \quad 16.8; \quad 16.4; \quad 17.1$$

1. Sulla base del campione precedente, testare $H_0 : \sigma^2 = 1.0$ usando un'ipotesi alternativa bilaterale e ad un livello $\alpha = 2\%$.
2. Che cosa possiamo concludere se la varianza campionaria fosse uguale a quella del punto 1., ma la dimensione del campione fosse $n = 25$?

### 1.5 Test su una proporzione

**Esercizio 13**

In 500 lanci di una moneta (truccata), è uscita testa 223 volte.

E' ragionevole "scommettere" che la probabilità $p$ che esca testa sia superiore al 40%? E al 50% ad un livello di significatività pari al 5%? Testare l'ipotesi nulla $p = 5$ ad un livello di significatività del 5%.

## 2 Svolgimenti

**Esercizio 1**

Indichiamo con $X$ la variabile aleatoria che rappresenta il calore emesso dal cemento. Per ipotesi, sappiamo che $X$ è distribuito come $N(\mu; 4)$.

1. L'errore di prima specie $\alpha$ è dato da:

$$\begin{aligned}
\alpha &= \mathbb{P}(\text{rigettare } H_0 \text{ dato che } H_0 \text{ è vera}) = \\
&= \mathbb{P}\left(\bar{X} < 98.5 \text{ oppure } \bar{X} > 101.5 \,\middle|\, \mu = 100\right) = \\
&= \mathbb{P}\left(\frac{\bar{X} - 100}{\frac{\sigma}{\sqrt{n}}} < \frac{98.5 - 100}{\frac{\sigma}{\sqrt{n}}}\right) + \mathbb{P}\left(\frac{\bar{X} - 100}{\frac{\sigma}{\sqrt{n}}} > \frac{101.5 - 100}{\frac{\sigma}{\sqrt{n}}}\right) \\
&= \phi(-2.25) + 1 - \phi(2.25) = 2(1 - \phi(2.25)) = 0.02445,
\end{aligned}$$

dove $\phi$ indica la funzione di ripartizione della normale standard.

2. L'errore di seconda specie $\beta$ è dato da:

$$\begin{aligned}
\beta &= \mathbb{P}\left(\text{accettare } H_0 \text{ dato che } H_0 \text{ è falsa}\right) = \\
&= \mathbb{P}\left(98.5 \leq \bar{X} \leq 101.5 \middle| \mu = 103\right) = \\
&= \mathbb{P}\left(\frac{98.5 - 103}{\frac{2}{3}} \leq \frac{\bar{X} - 103}{\frac{2}{3}} \leq \frac{101.5 - 103}{\frac{2}{3}}\right) = \\
&= \phi\left(-2.25\right) - \phi\left(-6.75\right) = \phi\left(6.75\right) - \phi\left(2.25\right) \simeq \\
&\simeq 1 - \phi\left(2.25\right) \simeq 0.012
\end{aligned}$$

La potenza del test è definita come $\left(1 - \beta\right)$, pertanto vale all'incirca 0.988.

3. In questo caso, $\beta$ è dato da:

$$\begin{aligned}
\beta &= \mathbb{P}\left(\text{accettare } H_0 \text{ dato che } H_0 \text{ è falsa}\right) = \\
&= \mathbb{P}\left(98.5 \leq \bar{X} \leq 101.5 \middle| \mu = 105\right) = \\
&= \mathbb{P}\left(\frac{98.5 - 105}{\frac{2}{3}} \leq \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \leq \frac{101.5 - 105}{\frac{2}{3}}\right) = \\
&= \phi\left(-5.25\right) - \phi\left(-9.75\right) \simeq 0
\end{aligned}$$

**Esercizio 2**

Indichiamo con $X$ la v.a. che rappresenta il diametro degli anelli dei pistoni. Per ipotesi $X$ è distribuita come $N\left(\mu; 10^{-6}\right)$.

La statistica test che si usa per questa verifica di ipotesi è $\bar{X}$ o, più precisamente, $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$. Per le ipotesi su $X$ ed applicando il TLC, è noto che la regione di accettazione del test

$$\begin{aligned}
H_0 &: \mu = 74.035 \\
\text{contro} \quad H_1 &: \mu \neq 74.035
\end{aligned}$$

ad un livello di significatività $\alpha$ è data da

$$-z_{\frac{\alpha}{2}} \leq \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq z_{\frac{\alpha}{2}},$$

dove $z_{\frac{\alpha}{2}}$ è definito (come in Montgomery) da $1 - \phi\left(z_{\frac{\alpha}{2}}\right) = \frac{\alpha}{2}$, ovvero l'area della coda alla destra di $z_{\frac{\alpha}{2}}$ è pari a $\frac{\alpha}{2}$.

1. Nel nostro caso, $z_{0.025} = 1.959964$ e la regione di accettazione al livello di significatività $\alpha = 5\%$ risulta essere:

$$\mu_0 - \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}} \leq \bar{X} \leq \mu_0 + \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}},$$

ovvero

$$74.03474 \leq \bar{X} \leq 74.03526$$

2. Siccome $\bar{x} = 74.036$ è al di fuori della regione di accettazione, allora non possiamo accettare l'ipotesi nulla (al livello $\alpha = 5\%$).

**Esercizio 3**

Indichiamo con $X$ la v.a. che rappresenta la vita della lampadina. Per ipotesi, sappiamo che $X$ è distribuita come $N(\mu; 625)$, che $n = 20$ e che $\bar{x} = 1014$.

1. In questo caso siamo interessati a verificare l'ipotesi alternativa del test

$$\begin{array}{rrcl} & H_0 & : & \mu \leq 1000 \\ \text{contro} & H_1 & : & \mu > 1000 \end{array}$$

   ad un livello di significatività $\alpha = 5\%$.

   E' noto che la regione di rifiuto per tale test è data da:

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > z_\alpha,$$

   dove $Z$ è approssimabile con una normale standard.

   Nel nostro caso $z_{0.05} = 1.644854$ e

$$\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{1014 - 1000}{25/\sqrt{20}} = 2.504396 > z_{0.05},$$

   pertanto rifiutiamo l'ipotesi nulla al livello $\alpha = 5\%$. E' quindi "ragionevole" supporre che la media di vita delle lampadine superi 1000 ore.

2. Ricordiamo che il $P - value$ è il più piccolo livello di significatività $\alpha$ che porta a rifiutare l'ipotesi $H_0$.

   Nel caso di un'ipotesi alternativa $H_1 : \mu > \mu_0$,

$$P - value = 1 - \phi\left(\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right),$$

   ovvero nel nostro caso $P - value = 1 - \phi\,(2.504396) = 0.006133$.

**Esercizio 4**

Indichiamo con $X$ la v.a. che rappresenta il tasso del processo. Per ipotesi, sappiamo che $\sigma^2 = 64$, che $n = 100$ e che $\bar{x} = 88.48\%$. Siccome la dimensione del campione è molto grande, è lecito utilizzare il TLC per calcolare la regione di accettazione del test anche se $X$ non è detto che sia normalmente distribuita.

1. In questo caso siamo interessati a verificare l'ipotesi

$$\begin{aligned} H_0 &: \mu = 90\% \\ \text{contro} \quad H_1 &: \mu \neq 90\% \end{aligned}$$

Il P-value del test è dato da

$$\begin{aligned} P - value &= 2\left(1 - \phi\left(\left|\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right|\right)\right) = \\ &= 0.05744 \end{aligned}$$

2. Siccome $\alpha = 5\%$ è inferiore al P-value, allora possiamo accettare $H_0$ al livello $\alpha = 5\%$.

   Siccome $\alpha = 6\%$ è superiore al P-value, allora dobbiamo rifiutare $H_0$ al livello $\alpha = 6\%$. Pertanto, ad un livello di significatività del 6% è ragionevole supporre che la media del tasso studiato non sia 90%.

**Esercizio 5**

Sappiamo che l'intervallo (bilaterale) di confidenza al $95\%(= 100\% - \alpha)$ per la media di una popolazione normale di varianza $\sigma^2$ e con un campione di dimensione $n$ è dato da:

$$\bar{x} - \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}} \leq \mu \leq \bar{x} + \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}.$$

Dai dati del problema, otteniamo che $\bar{x}$ e $n$ devono risolvere il seguente sistema:

$$\begin{cases} \bar{x} - \frac{\sqrt{6}}{\sqrt{n}} z_{0.025} = 11.4104 \\ \bar{x} + \frac{\sqrt{6}}{\sqrt{n}} z_{0.025} = 13.8896 \end{cases}$$

ovvero

$$\begin{cases} \bar{x} = 12.65 \\ n = 15 \end{cases}$$

**Esercizio 6**

Dai 10 dati a disposizione sulla lunghezza delle mine, otteniamo una media campionaria pari a $\bar{x} = 12.933$.

1. L'intervallo bilaterale di confidenza al 95% per la media è dato quindi da

$$\bar{x} - \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}} \leq \mu \leq \bar{x} + \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}},$$

   ovvero da

$$\mu \in [12.871; 12.995]$$

2. Osserviamo innanzitutto che l'ampiezza dell'intervallo di confidenza per $\mu$ al livello $\alpha$ è pari a $\left[\left(\bar{x} + \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}\right) - \left(\bar{x} - \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}\right)\right] = 2\frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}$. Per avere un intervallo non più ampio di $10^{-2}$, è necessario e sufficiente che

$$\begin{aligned} 2\frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}} &\leq 10^{-2} \\ \sqrt{n} &\geq 200\sigma z_{0.025} \\ n &\geq 1536.584 \end{aligned}$$

   quindi la dimensione minima del campione è pari a $n = 1537$.

**Esercizio 7**

Indichiamo con $X$ la v.a. che rappresenta il contenuto delle bottiglie di champagne. Per ipotesi $X$ è distribuita come $N\left(\mu; \sigma^2\right)$, dove entrambi i parametri non sono noti. Indichiamo con $\bar{X}$ la media campionaria e con $S^2$ la varianza campionaria.

La statistica test che si usa per questa verifica di ipotesi è $\frac{\bar{X}-\mu}{S/\sqrt{n}}$ che, sotto ipotesi di normalità di $X$ o di grandi dimensioni del campione, è distribuita come una $t\,(n-1)$, ovvero una $t - \textit{student}$ con $(n-1)$ gradi di libertà. E' anche noto che la regione di rifiuto del test

$$\begin{aligned} H_0 &: \mu \geq 730 \\ \text{contro} \quad H_1 &: \mu < 730 \end{aligned}$$

ad un livello di significatività $\alpha$ è data da

$$T \triangleq \frac{\bar{X} - \mu_0}{S/\sqrt{n}} < -t_{\alpha,n-1},$$

dove $t_{\alpha,n-1}$ è definito (come in Montgomery) da $\mathbb{P}\left(T > t_{\alpha,n-1}\right) = \alpha$, ovvero l'area della coda alla destra di $t_{\alpha,n-1}$ è pari a $\alpha$.

Nel nostro caso, quindi, la regione di rifiuto al livello di significatività $\alpha = 5\%$ è data da:

$$\bar{X} < \mu_0 - \frac{S}{\sqrt{n}} t_{\alpha,n-1} = 725.3774$$

Siccome la nedia campionaria a noi disponibile è $\bar{x} = 726$, non possiamo rifiutare l'ipotesi nulla, quindi non possiamo concludere che le bottiglie contengano meno di quanto dichiarato.

**Esercizio 8**

Indichiamo con $X$ l'altezza delle reclute nel 1980. Vogliamo testare l'ipotesi

$$H_0 : \mu = 169$$

contro

$$H_1 : \mu \neq 169$$

La regione di accettazione dell'ipotesi nulla al livello di significatività $\alpha = 5\%$ è data da

$$-1.9799 = -t_{0.025;120} \leq \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \leq t_{0.025;120} = 1.9799.$$

Siccome, nel nostro caso,

$$\frac{\bar{x} - \mu_0}{S/\sqrt{n}} \simeq 2.386$$

allora dobbiamo rifiutare l'ipotesi nulla, quindi non possiamo affermare che l'altezza delle reclute sia rimasta invariata.

**Esercizio 9**

Indichiamo con $X$ la v.a che rappresenta la resistenza degli elementi prodotti.

1. L'intervallo bilaterale di confidenza per la media al $95\% = 100\% - \alpha$ nel caso di media e varianza incognite è dato da

$$\bar{X} - t_{\frac{\alpha}{2};n-1} \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{\frac{\alpha}{2};n-1} \frac{S}{\sqrt{n}}$$

   che nel nostro caso diventa

$$\mu \in [1.06; 1.34] .$$

2. Siccome $\mu_0 = 1.2$ appartiene all'intervallo bilaterale che abbiamo appena trovato, allora possiamo accettare l'ipotesi nulla al livello $\alpha = 5\%$.

**Esercizio 10**

Indichiamo con $X$ la v.a che rappresenta la concentrazione dell'enzima nel sangue.

1. Ragionando come nell'esercizio precedente, otteniamo che, nel caso di media e varianza incognite, l'intervallo bilaterale al 95% per la media è dato da:
$$\mu \in [0.934; 1.526]$$

2. Nel caso di varianza nota pari a $\sigma^2 = 0.4$, si ottiene, invece, che l'intervallo di confidenza diventa
$$\mu \in [0.953; 1.507]$$

3. Ovviamente, è più ampio l'intervallo trovato in 1., dove la varianza è incognita e l'incertezza è maggiore.

**Esercizio 11**

Indichiamo con $X$ la v.a. che rappresenta il contenuto di zucchero nei succhi. Per ipotesi $X$ è distribuita come $N\left(\mu; \sigma^2\right)$, dove entrambi i parametri non sono noti. Indichiamo con $S^2$ la varianza campionaria.

La statistica test che si usa per una verifica di ipotesi sulla varianza è $\frac{(n-1)S^2}{\sigma^2}$ che, sotto ipotesi di normalità di $X$ o di grandi dimensioni del campione, è distribuita come una $\chi^2(n-1)$, ovvero una $chi - quadro$ con $(n-1)$ gradi di libertà. E' anche noto che la regione di accettazione del test

$$\begin{array}{rrcl} & H_0 & : & \sigma^2 = 18 = \sigma_0^2 \\ \text{contro} & H_1 & : & \sigma^2 \neq 18 \end{array}$$

ad un livello di significatività $\alpha$ è data da

$$\chi^2_{1-\frac{\alpha}{2},n-1} \leq \frac{(n-1)\,S^2}{\sigma_0^2} \leq \chi^2_{\frac{\alpha}{2},n-1},$$

dove $\chi^2_{\frac{\alpha}{2},n-1}$ è definito (come in Montgomery) da $\mathbb{P}\left(\cdot > \chi^2_{\frac{\alpha}{2},n-1}\right) = \frac{\alpha}{2}$, ovvero l'area della coda alla destra di $\chi^2_{\frac{\alpha}{2},n-1}$ è pari a $\alpha/2$.

1. Nel nostro caso, quindi, la regione di accettazione al livello di significatività $\alpha = 5\%$ è data da:
$$\frac{(n-1)\,S^2}{\sigma_0^2} \in [2.7004; 19.023]$$

   Siccome $\frac{(n-1)s^2}{\sigma_0^2} = 11.57 \notin$regione di accettazione, allora accettiamo l'ipotesi nulla al livello del 5%.

2. L'intervallo di confidenza al 98% per la varianza è dato da

$$\frac{(n-1)\,S^2}{\chi^2_{\frac{\alpha}{2},n-1}} \leq \sigma^2 \leq \frac{(n-1)\,S^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}$$

che, nel nostro caso, corrisponde a

$$\sigma^2 \in [9.57; 99.32]$$

3. Con l'aiuto del pto precedente, possiamo accettare l'ipotesi nulla al livello di significatività $\alpha = 2\%$, siccome $\sigma_0^2 = 18$ appartiene all'intervallo di confidenza trovato sopra.

**Esercizio 12**

Indichiamo con $X$ la v.a. che rappresenta il livello di grassi insaturi presenti nella margarina.

Vogliamo testare la seguente ipotesi nulla sulla varianza di $X$

$$H_0 : \sigma^2 = 1.0 = \sigma_0^2$$

contro l'ipotesi alternativa

$$H_1 : \sigma^2 \neq 1.0$$

Supponendo che $X$ sia normalmente distribuita, la regione di accettazione (al livello di significatività $\alpha$) per il test precedente è data da

$$\chi^2_{1-\frac{\alpha}{2},n-1} \leq \frac{(n-1)\,S^2}{\sigma_0^2} \leq \chi^2_{\frac{\alpha}{2},n-1}$$

Dai dati a noi noti, inoltre, ricaviamo che

$$\begin{aligned}
\bar{x} &= \frac{1}{6}\,[16.6 + 17.1 + 17.4 + 16.8 + 16.4 + 17.1] = 16.9 \\
s^2 &= \frac{1}{5}\sum_{i=1}^{6}(x_i - \bar{x})^2 = \\
&= \frac{1}{5}\left[(-0.3)^2 + (0.2)^2 + (0.5)^2 + (-0.1)^2 + (-0.5)^2 + (0.2)^2\right] = 0.136
\end{aligned}$$

1. Osserviamo che la regione di accettazione dell'ipotesi nulla al livello di significatività $\alpha = 2\%$ è data da

$$\frac{(n-1)\,S^2}{\sigma_0^2} \in [0.5543; 15.086]$$

che contiene $\frac{(n-1)s^2}{\sigma_0^2} = 0.68$. Quindi possiamo accettare l'ipotesi nulla!

2. Se il livello di significatività $\alpha$ e la varianza campionaria rimanessero uguali a quelle del punto precedente, ma $n = 25$, la regione di accettazione di $H_0$ diventerebbe:
$$\frac{(n-1)\,S^2}{\sigma_0^2} \in [10.8564; 42.980]\,.$$
In questo caso, $\frac{(n-1)s^2}{\sigma_0^2} = 24 \cdot 0.136 = 3.264$ non appartiene ala regione di accettazione, quindi dobbiamo rifiutare l'ipotesi nulla!

**Esercizio 13**

Indichiamo con $X$ la v.a. che conta il numero di volte che esce testa. E' chiaro quindi che $X$ è distribuita come una binomiale con percentuale di successo data da $p$.

Indichiamo con $\hat{p} \triangleq \frac{X}{n}$ la proporzione di teste in $n$ lanci della moneta. $\hat{p}$, quindi, è uno stimatore di $p$.

La statistica test che si usa per una verifica di ipotesi su una proporzione è $\frac{X-np}{\sqrt{np(1-p)}}$ che, nel caso in cui $np > 5$ e $n\,(1-p) > 5$, è approssimabile con una normale standard.

Sia
$$H_0 : p \geq p_0 = 0.4$$
contro l'ipotesi alternativa
$$H_1 : p < p_0 = 0.4$$
e calcoliamo il P-value del test come
$$\overline{\alpha} = \phi\left(\frac{\overline{x}_n - p_0}{\sqrt{p_0(1-p_0)/n}}\right) = \phi\left(\frac{\overline{0}.466 - 0.4}{\sqrt{0.4 \cdot 0.6/500}}\right) = \phi(3.0125) = 0.9987.$$
Poichè il $P$-value è molto vicino a 1, si accetta l'ipotesi nulla. Se ora l'ipotesi nulla è
$$H_0 : p \geq p_0 = 0.5$$
contro l'ipotesi alternativa
$$H_1 : p < p_0 = 0.5$$
e calcoliamo il P-value del test come
$$\overline{\alpha} = \phi\left(\frac{\overline{x}_n - p_0}{\sqrt{p_0(1-p_0)/n}}\right) = \phi\left(\frac{\overline{0}.466 - 0.5}{\sqrt{0.5 \cdot 0.5/500}}\right) = \phi(-1.5205) = 0.0642.$$
Ad un livello di significatività del 5% l'ipotesi nulla è accettata. Consideriamo infine il seguente test
$$H_0 : p = p_0 = 0.5$$

contro l'ipotesi alternativa

$$H_1 : p \neq p_0 = 0.5$$

il cui P-value è

$$\overline{\alpha} = 2-2\phi\left(\left|\frac{\overline{x}_n - p_0}{\sqrt{p_0(1-p_0)/n}}\right|\right) = \phi\left(\left|\frac{\overline{0}.466 - 0.5}{\sqrt{0.5 \cdot 0.5/500}}\right|\right) = 2-2\phi(1.5205) = 0.1284.$$

Ad un livello di significatività del 5% accettiamo nuovamente l'ipotesi nulla.