Esame di *Calcolo delle probabilità e statistica* (per studenti di Informatica)
corso A
Università degli studi di Bari Aldo Moro

09-06-2020

Esercizio 1. Una scatola contiene 100 monete, 10 delle quali sono truccate e danno testa con probabilità $p = 0.35$, e le restanti 90 sono equilibrate.

(1) Si estraggono *insieme* due monete, che poi vengono lanciate. Sapendo che hanno entrambe dato testa, qual è la probabilità che siano entrambe truccate?
(2) Quante volte n è necessario lanciare una moneta truccata perché la probabilità di avere un numero di teste compreso tra $0.30n$ e $0.40n$ sia almeno del 90%? (Usare l'approssimazione normale data dal teorema del limite centrale.)

Esercizio 2. Si verifichi che la funzione $f(x) = \theta a^\theta x^{-(\theta+1)}$ per $x \geq a$ e $f(x) = 0$ per $x < a$, dove $a > 0$ e $\theta > 1$ sono due parametri, è la densità di probabilità di una certa variabile aleatoria X. Supponendo che il parametro a sia noto:

(1) determinare lo stimatore di massima verosimiglianza per θ corrispondente a un campione $(X_1, X_2, \ldots, X_n)$ di rango n.
(2) calcolare l'aspettazione $E[X]$ della variabile X che ha f come densità e trovare uno stimatore corretto per la quantità $\frac{\theta}{\theta-1}$.

Esercizio 3. Un campione di 16 sigarette di una certa marca è stato sottosposto ad analisi per misurarne il contenuto di nicotina. Per la media e la varianza campionaria si sono ottenuti i seguenti risultati: $\bar{x} = 20\, mg$ e $S^2 = 4\,(mg)^2$.

(1) Calcolare un intervallo (bilaterale) di confidenza al 95% per il contenuto medio di nicotina X di ogni sigaretta, supponendo che $X \sim N(\mu, \sigma^2)$, con μ e σ^2 parametri incogniti.
(2) Calcolare un intervallo destro (cioè una semiretta del tipo $[\sigma^2_{min}, +\infty)$) di confidenza al 95% per σ^2.

Esercizio 1

1. H_1 = "Escono due monete truccate"
   H_2 = "Escono una moneta truccata e una equa"
   H_3 = "Escono due monete eque"

   E = "Escono due teste"

Dobbiamo calcolare $p(H_1|E)$, che è data dalla formula di Bayes

$$p(H_1|E) = \frac{p(E|H_1)\cdot p(H_1)}{\sum_{i=1}^{3} p(E|H_i)\,p(H_i)}$$

Ora $p(H_1) = \dfrac{\binom{10}{2}}{\binom{100}{2}} = \dfrac{\frac{10!}{2!\,8!}}{\frac{100!}{2!\,98!}} = \dfrac{\frac{10\cdot 9}{2}}{\frac{100\cdot 99}{2}} = \dfrac{10\cdot 9}{100\cdot 99} = \dfrac{1}{10}\cdot\dfrac{1}{11}$

$$p(H_2) = \frac{\frac{90\cdot 10}{\cancel{2}}}{\binom{100}{2}} = \frac{90\cdot 10}{\frac{100\cdot 99}{\cancel{2}}} = 2\left(\frac{90}{100}\cdot\frac{10}{99}\right) = 2\left(\frac{\cancel{9}}{\cancel{10}}\cdot\frac{\cancel{10}}{\cancel{99}_{11}}\right) = \frac{2}{11}$$

$$p(H_3) = \frac{\binom{90}{2}}{\binom{100}{2}} = \frac{\frac{90!}{2!\,88!}}{\frac{100!}{2!\,98!}} = \frac{\frac{90\cdot 89}{2}}{\frac{100\cdot 99}{2}} = \frac{90}{100}\cdot\frac{89}{99} =$$

$$\frac{\cancel{9}}{10}\cdot\frac{89}{\cancel{99}_{11}} = \frac{89}{10\cdot 11}$$

N.B. $p(H_1) + p(H_2) + p(H_3) = 1$

$p(E|H_1) = 0.35^2$ e $p(E|H_3) = 0.5^2$

$p(E|H_2) = 0.5 \times 0.35$

Inserendo i dati numerici troviamo

$$p(H_1|E) = \frac{0.35^2 \; \frac{1}{10} \cdot \frac{1}{11}}{0.35^2 \; \frac{1}{10} \cdot \frac{1}{11} + 0.5 \times 0.35 \cdot \frac{2}{11} + 0.5^2 \; \frac{81}{10 \cdot 11}}$$

$$= \frac{0.01225}{0.01225 + 0.0318 + 2.225} \simeq 5.4 \cdot 10^{-3}$$

(molto bassa)

2. Sia X_i, $i = 1, 2, \ldots, n$, la variabile bernoulliana che vale 1 con probabilità $p = 0.35$ e 0 con probabilità $1-p = 0.35$

Vogliamo che

$$P\left(0.30n \le \sum_{i=1}^{n} X_i \le 0.40n\right) \ge 0.90$$

Ora $P\left(0.30n \le \sum_{i=1}^{n} X_i \le 0.40n\right) =$

$$P\left(\frac{0.30n - 0.35n}{\sqrt{0.35 \times 0.65\, n}} \le \frac{\sum_{i=1}^{n} X_i - 0.35n}{\sqrt{0.35 \times 0.65\, n}} \le \frac{0.40n - 0.35n}{\sqrt{0.35 \times 0.65\, n}}\right)$$

$$\uparrow$$

$$Z \sim N(0,1)$$

Quindi $P\left(|Z| \le \frac{0.05 \sqrt{n}}{0.4769696}\right) =$

$$P\left(|Z| \le 0.1048\sqrt{n}\right) =$$

$$Erf(0.1048\sqrt{n}) - Erf(-0.1048\sqrt{n}) =$$

$$Erf(0.1048\sqrt{n}) - (1 - Erf(0.1048\sqrt{n})) =$$

$$2Erf(0.1048\sqrt{n}) - 1 \overset{!}{=} 0.90$$

$$Erf(0.1048\sqrt{n}) = 0.95$$

$$\Rightarrow \quad 0.1048\sqrt{n} = 1.65 \quad \Rightarrow n \ge 248.$$

Esercizio 2

$$f(x) = \begin{cases} \theta a^{\theta} x^{-(\theta+1)} & x \geq a \\ 0 & x < a \end{cases}$$

$f(x) \geq 0 \quad \forall x \in \mathbb{R}$ e $\int_{-\infty}^{\infty} f(x)dx = 1$. Infatti, abbiamo

$$\int_{-\infty}^{\infty} f(x)dx = \int_{a}^{+\infty} \theta a^{\theta} x^{-\theta-1} dx = \theta a^{\theta} \left. \frac{x^{-\theta}}{-\theta} \right|_{a}^{\infty} = \cancel{\theta} a^{\theta} \frac{1}{\cancel{\theta}} a^{-\theta} = 1.$$

1. $f(x_1, x_2, \dots, x_n | \theta) = \theta^n a^{n\theta} (x_1 \dots x_n)^{-(\theta+1)}$ per $x_1 \geq a$, $x_2 \geq a$, $\dots$, $x_n \geq a$

$$\ln f(x_1 \dots x_n | \theta) = n \ln \theta + n\theta \ln a - (\theta+1) \ln(x_1 \dots x_n)$$

$$\frac{\partial}{\partial \theta} \ln f(x_1 \dots x_n | \theta) = \frac{n}{\theta} + n \ln a - \ln(x_1 x_2 \dots x_n) \stackrel{!}{=} 0$$

$$\frac{n}{\theta} = \ln(x_1 \dots x_n) - n \ln a$$

$$\frac{\theta}{n} = \frac{1}{\ln(x_1 x_2 \dots x_n) - n \ln a}; \quad \theta = \frac{1}{\frac{1}{n} \sum_{i=1}^{n} \ln x_i - \ln a}$$

$$\hat{\theta}_n = \frac{1}{\frac{1}{n} \sum_{i=1}^{n} \ln X_i - \ln a}$$

2. $$E[X] = \int_{a}^{+\infty} x \, \theta a^{\theta} x^{-(\theta+1)} dx = \theta a^{\theta} \int_{a}^{\infty} x^{-\theta} dx =$$

$$= \theta \left. \frac{a^{\theta} x^{-\theta+1}}{-\theta+1} \right|_{a}^{\infty} =$$

$$= \theta \frac{a^{\theta} a^{-\theta+1}}{1-\theta} = \frac{\theta}{1-\theta} a$$

Quindi $\frac{1}{a} E[X] = \frac{\theta}{1-\theta}$

per cui $\frac{1}{a}\overline{X} = \frac{1}{a}\frac{1}{n}\sum_{i=1}^{n} X_i$ è uno stimatore corretto di $\frac{\theta}{1-\theta}$

Esercizio 3

1. $n=16$

$\overline{X} = 20\ mg \quad S^2 = 4\ (mg)^2 \Rightarrow S = 2\ mg$

$$\left[\overline{X} - \frac{S}{\sqrt{n}}t_{1-\frac{\alpha}{2}}(n-1);\ \overline{X} + \frac{S}{\sqrt{n}}t_{1-\frac{\alpha}{2}}(n-1)\right]$$

$1-\alpha = 0.95;\quad \alpha = 0.05 \Rightarrow 1-\frac{\alpha}{2} = 0.975$

$$t_{0.975}(15) = 2.1315$$

$$\frac{S}{\sqrt{n}}t_{1-\frac{\alpha}{2}}(n-1) = \frac{2}{4}\,2.1315 = \frac{2.135}{2} = 1.06\ 75$$

$\Rightarrow [18.93;\ 21.07]$

2. Da $P\left[\frac{(n-1)S^2}{\sigma^2} \leq \chi^2_{1-\alpha}(n-1)\right] = 1-\alpha$

Vediamo che $P\left[\frac{\sigma^2}{(n-1)S^2} \geq \frac{1}{\chi^2_{1-\alpha}(n-1)}\right] = 1-\alpha$

quindi $\left[\frac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)},\ +\infty\right)$ è l'intervallo richiesto.

Con i nostri dati

$$\frac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)} = \frac{(15)\cdot 4}{\chi^2_{0.95}(15)} = \frac{15\cdot 4}{24.996} = 2.40$$

L'intervallo è dunque $[2.40,\ +\infty)$

Esame di *Calcolo delle probabilità e statistica* (per studenti di Informatica)
corso A
Università degli studi di Bari Aldo Moro

26-06-2020

Esercizio 1. Si lancia una moneta equilibrata. Se esce testa, si lancia $n \geq 1$ volte un dado equo a 6 facce numerate da 1 a 6; se esce croce, lo si lancia $2n$ volte.

(1) Calcolare (in funzione di n) la probabilità che 1 esca almeno una volta.
(2) Calcolare il limite per $n \to \infty$ della probabilità trovata sopra.
(3) Calcolare la probabilità che la moneta abbia dato testa sapendo che 1 è uscito esattamente n volte.

Esercizio 2.
Verificare che la funzione $f(x) = \theta^2 x e^{-\theta x}$ per $x \geq 0$ e nulla altrimenti è la densità di probabilità di una certa variabile aleatoria X per ogni valore del parametro $\theta > 0$.

(1) Determinare lo stimatore di massima verosimiglianza per il parametro θ corrispondente a un campione $(X_1, X_2, \ldots, X_n)$ di rango n.
(2) Calcolare l'aspettazione $E[X]$ della variabile aleatoria X e dire se lo stimatore ottenuto al punto precedente è consistente.

Esercizio 3. Una macchina è impostata in modo che il contenuto medio di succo per bottiglia sia uguale a μ. Un campione di 100 bottiglie fornisce un contenuto medio di 48.8 cl. Assumendo che la deviazione standard della popolazione sia nota e pari a 5 cl:

(1) Verificare l'ipotesi che il contenuto medio per bottiglia sia 50 cl a un livello di significatività del 10% e del 5%.
(2) Calcolare il p-value del test.

Esercizio 1

T = "Esce testa" $\quad p(T)=\frac{1}{2}$

C = "Esce croce" $\quad p(C)=\frac{1}{2}$

E = "Esce almeno una volta 1"

1. $p(E) = 1 - p(E^c)$ $\qquad E^c$ = "Non esce <u>mai</u> 1"

Ora

$$p(E^c) = p(E^c|T)\cdot p(T) + p(E^c|C)\cdot p(C)$$ (formula della probabilità totale)

$p(E^c|T)$ è la probabilità che in n lanci di ~~un~~ dado equilibrato non esca mai 1, perciò

$$p(E^c|T) = \left(\frac{5}{6}\right)^n$$

Analogamente

$$p(E^c|C) = \left(\frac{5}{6}\right)^{2n}$$

$$p(E^c) = \left(\frac{5}{6}\right)^n\cdot\frac{1}{2} + \left(\frac{5}{6}\right)^{2n}\cdot\frac{1}{2} = \frac{1}{2}\left(\frac{5}{6}\right)^n\left[1+\left(\frac{5}{6}\right)^n\right]$$

$$p(E) = 1 - \frac{1}{2}\left(\frac{5}{6}\right)^n\left[1+\left(\frac{5}{6}\right)^n\right]$$

2. $\lim_{n\to\infty} p(E) = 1$ perché $\left(\frac{5}{6}\right)^n \xrightarrow[n\to\infty]{} 0$

3. $p(T \mid E_1)$

$E_1 =$ " il numero 1 è uscito n-volte"

$$p(T|E_1) = \frac{p(E_1|T)\cdot p(T)}{p(E_1|T)\cdot p(T) + p(E_1|C)\cdot p(C)}$$

$$p(E_1|T) = \left(\frac{1}{6}\right)^n$$

$$p(E_1|C) = \binom{2n}{n}\left(\frac{1}{6}\right)^n\left(\frac{5}{6}\right)^n = \frac{(2n)!}{n!\,n!}\left(\frac{1}{6}\right)^n\left(\frac{5}{6}\right)^n$$

$$p(T|E_1) = \frac{\cancel{\frac{1}{2}}\ \cancel{\left(\frac{1}{6}\right)^n}}{\cancel{\frac{1}{2}}\ \cancel{\left(\frac{1}{6}\right)^n} + \frac{(2n)!}{n!\,n!}\cancel{\left(\frac{1}{6}\right)^n}\left(\frac{5}{6}\right)^n \cancel{\frac{1}{2}}} = \frac{1}{1 + \frac{(2n)!}{(n!)^2}\left(\frac{5}{6}\right)^n}$$

Esercizio 2

$$f(x) = \begin{cases} \theta^2 x\, e^{-\theta x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Si ha $f(x) \geq 0 \ \forall x \in \mathbb{R}$. Inoltre, $\int_{-\infty}^{\infty} f(x)dx = 1$. Infatti,

$$\int_{-\infty}^{\infty} f(x)dx = \int_0^{\infty} \theta^2 x\, e^{-\theta x} dx = \theta^2 \int_0^{\infty} \underbrace{x}_{g}\, \underbrace{e^{-\theta x}}_{f'} dx = \theta^2 \left[\frac{e^{-\theta x}}{-\theta} x \Big|_0^{\infty} - \int_0^{\infty} \frac{e^{-\theta x}}{-\theta} dx \right]$$

$$= \theta \int_0^{\infty} e^{-\theta x} dx = \theta \left. \frac{e^{-\theta x}}{-\theta} \right|_0^{\infty} = 1.$$

1) $f(x_1, x_2, \dots, x_n | \theta) = \theta^{2n} x_1 x_2 \dots x_n\, e^{-\theta(x_1 + x_2 + \dots + x_n)}$

$L(x_1, x_2, \dots, x_n|\theta) = \ln f(x_1, x_2, \dots, x_n|\theta) =$

$$2n \ln\theta + \ln(x_1 \dots x_n) - \theta(x_1 + x_2 + \dots + x_n)$$

$$\frac{\partial L}{\partial \theta} = \frac{2n}{\theta} - (x_1 + x_2 + \dots + x_n) \stackrel{!}{=} 0$$

$$\frac{2n}{\theta} = x_1 + x_2 + \dots + x_n ; \quad \frac{\theta}{2n} = \frac{1}{x_1 + x_2 + \dots + x_n} ; \quad \theta = \frac{2}{\frac{x_1 + x_2 + \dots + x_n}{n}}$$

$\hat{\theta}_n = \dfrac{2}{\overline{X}_n}$ con $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ (media campionaria)

2) $$E[X] = \int_{-\infty}^{\infty} x f(x) dx = \int_0^{\infty} \theta^2 x^2 e^{-\theta x} dx = \theta^2 \int_0^{\infty} \underbrace{x^2}_{g} \underbrace{e^{-\theta x}}_{f'} dx =$$

$$\theta^2 \left[\frac{e^{-\theta x}}{-\theta} x^2 \Big|_0^{\infty} - \int_0^{\infty} \frac{e^{-\theta x}}{-\theta} \cdot 2x\, dx \right] = 2\theta \int_0^{\infty} x\, e^{-\theta x} dx$$

$$= \frac{2\theta}{\theta^2} = \frac{2}{\theta}$$

Quindi $E[X] = \frac{2}{\theta}$

$\hat{\theta}_n$ è consistente se $\hat{\theta}_n$ converge in probabilità a θ.

Ora dalla legge debole dei grandi numeri

$$\overline{X}_n \xrightarrow{P} E[X] = \frac{2}{\theta}$$

per cui

$$\frac{2}{\overline{X}_n} \xrightarrow{P} \theta$$

cioè

$$\hat{\theta}_n \xrightarrow{P} \theta.$$

3

$n = 100$

$\overline{x} = 48.8 \text{ cl}$

$\sigma = 5 \text{ cl}$

$H_0: \mu = \mu_0$ con $\mu_0 = 50$

$\alpha = 0.1$ e $\alpha = 0.05$

Dato che la varianza è nota, la statistica da usare è

$\frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ che è distribuita come una normale standard

Si accetta H_0 se $\left| \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \right| \leq z_{1-\frac{\alpha}{2}}$

Ora $\left|\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| = \frac{|48.8-50|}{5} \cdot 10 = 1.2 \cdot 2 = 2.4$

- $\alpha = 0.1$ $\quad \frac{\alpha}{2} = 0.05 \quad 1-\frac{\alpha}{2} = 0.95$ e

$z_{0.95} = 1.65$

Quindi H_0 viene rifiutata a un livello di significatività del 10%. (perché $2.4 > 1.65$)

$\alpha = 0.05$ $\quad \frac{\alpha}{2} = 0.025 \quad 1-\frac{\alpha}{2} = 0.975$ e

$z_{0.975} = 1,96$

Quindi H_0 viene rifiutata anche a un livello di significatività del 5%. (perché $2.4 > 1.96$)

Il p-value è il più piccolo livello di significatività al di sopra del quale si rifiuta H_0.
E' dunque determinato dalla condizione

$$z_{1-\frac{p}{2}} = 2.4$$

così $1-\frac{p}{2} = Erf(2.4) \Rightarrow p = 2(1-Erf(2.4)) =$
$2(1-0.9918) = 0.0164$

quindi il p-value è circa 1%. (c'è evidenza sperimentale contro H_0)

Esame di *Calcolo delle probabilità e statistica* (per studenti di Informatica)

corso A

Università degli studi di Bari Aldo Moro

17-07-2020

Esercizio 1.
Si lancia una moneta equilibrata N volte, dove N è una variabile aleatoria di Poisson di parametro $\lambda = 2$.

(1) Calcolare la probabilità che esca almeno una testa.
(2) Calcolare la probabilità che N sia pari a n sapendo che non è mai uscita testa verificando che è pari a $\frac{1}{(e-1)n!}$.
(3) Calcolare il limite per $n \to \infty$ della probabilità determinata al punto precedente.

(Per svolgere i calcoli è utile ricordare che $\sum_{k=0}^{\infty} \frac{1}{k!} = e$ e quindi $\sum_{k=1}^{\infty} \frac{1}{k!} = e - 1$)

Esercizio 2. Verificare che per ogni $\theta > 0$ la funzione $f(x) := 2\theta x e^{-\theta x^2}$ per $x \geq 0$ e nulla altrimenti è la densità di probabilità di una certa variabile aleatoria X.

(1) Determinare lo stimatore di massima verosimiglianza per θ corrispondente a un campione $(X_1, X_2, \cdots, X_n)$ di rango n distribuito come X.
(2) Esibire una statistica sufficiente per il campione $(X_1, X_2, ..., X_n)$ (sfruttando il teorema di fattorizzazione di Fisher).

Esercizio 3.
Un primo campione di rango $n_1 = 9$ segue legge gaussiana con varianza nota pari a $\sigma_1^2 = 1.32$ e media μ_1 incognita; un secondo campione di rango $n_2 = 16$ segue anch'esso legge gaussiana con varianza nota pari a $\sigma_2^2 = 2.15$ e media μ_2 incognita. Si osservano medie campionarie date rispettivamente da $\overline{\mu_1} = 10.20$ e $\overline{\mu_2} = 11.15$.

(1) Condurre un test di verifica dell'ipotesi $\mu_1 = \mu_2$ a un livello di significatività del 10% e del 5%.
(2) Calcolare il p-value del test considerato sopra.

Esercizio n°1

Dato che N è una v.a. di Poisson, si ha

$$P(N=k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

Nel nostro caso $\lambda = 2$, perciò $P(N=k) = e^{-2} \frac{2^k}{k!}$

1) E = "Esce almeno una testa"

$$p(E) = 1 - p(E^c)$$

con

E^c = "Non esce nemmeno una testa"

Ora

$$p(E^c) = \sum_{k=1}^{\infty} p(E^c | N=k) \cdot p(N=k)$$ (formula della probabilità totale)

$p(E^c | N=k)$ è la probabilità che in k lanci esca sempre croce. Pertanto abbiamo

$$P(E^c | N=k) = \left(\frac{1}{2}\right)^k$$

In definitiva

$$P(E^c) = \sum_{k=1}^{\infty} \left(\frac{1}{2}\right)^k e^{-2} \frac{2^k}{k!} = e^{-2} \sum_{k=1}^{\infty} \frac{1}{k!} = \frac{1}{e^2}(e-1) = \frac{e-1}{e^2}$$

$$P(E) = 1 - \frac{e-1}{e^2} = \frac{e^2-e+1}{e^2} \simeq 0.77$$

2) $P(N=n \mid E^c) = \dfrac{p(E^c \mid N=n) \cdot p(N=n)}{p(E^c)}$

$$= \frac{\left(\frac{1}{2}\right)^n \cdot e^{-2} \frac{2^n}{n!}}{\frac{e-1}{e^2}} = \frac{e^{-2}}{n!} \frac{e^2}{e-1} = \frac{1}{e-1} \frac{1}{n!}$$

3) $\lim_{n \to \infty} P(N=n \mid E^c) = \lim_{n \to \infty} \frac{1}{(e-1)} \frac{1}{n!} = 0$

(perché $n! \to +\infty$)

Esercizio n° 2

$$f(x) = \begin{cases} 2\theta x e^{-\theta x^2} & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases}$$

i) $f(x) \geq 0 \quad \forall x \in \mathbb{R}$

e

ii $\int_{\mathbb{R}} f(x)\,dx = 1$. Infatti, $\int_{\mathbb{R}} f(x)\,dx = \int_0^\infty 2\theta x \, e^{-\theta x^2}\,dx$

$$= -e^{-\theta x^2} \Big|_0^\infty = 1$$

1) $f(x_1, x_2, \ldots, x_n \mid \theta) = 2\theta x_1 e^{-\theta x_1^2} \cdot 2\theta x_2 e^{-\theta x_2^2} \cdot \ldots$

$$2\theta x_n e^{-\theta x_n^2} =$$

$$2^n \theta^n x_1 x_2 \ldots x_n \, e^{-\theta \sum_{i=1}^n x_i^2}$$

$\ln f(x_1, x_2, \ldots, x_n \mid \theta) = n \log \theta - \theta \sum_{i=1}^n x_i^2$ (a meno di costanti additive che non dipendono da θ)

$$\frac{\partial}{\partial\theta}\ln f(x_1,\dots,x_n|\theta) = \frac{n}{\theta} - \sum_{i=1}^{n} x_i^2 \stackrel{!}{=} 0$$

$$\frac{n}{\theta} = \sum_{i=1}^{n} x_i^2 \Rightarrow \frac{\theta}{n} = \frac{1}{\sum_{i=1}^{n} x_i^2} \Rightarrow \theta = \frac{n}{\sum_{i=1}^{n} x_i^2}$$

$\hat{\theta}_n = \frac{n}{\sum_{i=1}^{n} X_i^2}$ è lo stimatore di massima verosimiglianza

② $f(x_1,\dots,x_n|\theta) = \underbrace{2^n x_1 x_2 \dots x_n}_{u(x)} \underbrace{\theta^n e^{-\theta\sum_{i=1}^{n} x_i^2}}_{v(\theta, t(x))}$

con $t(x) = \sum_{i=1}^{n} x_i^2$

Ne segue che $T = \sum_{i=1}^{n} X_i^2$ è una statistica sufficiente.

Esercizio 3

$n_1 = 9$

$n_2 = 16$

$\sigma_1^2 = 1{,}32$

$\sigma_2^2 = 2{,}15$

$\overline{\mu}_1 = 10{,}20$

$\overline{\mu}_2 = 11{,}15$

① $H_0: \mu_1 = \mu_2$ con $\alpha = 0.1$ e poi con $\alpha = 0.05$

La statistica da utilizzare per condurre il test è

$$\frac{(\overline{X}-\mu_1)-(\overline{Y}-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$$

che è distribuita
come una normale
standard

Sotto H_0 $\mu_1 = \mu_2$, perciò abbiamo la statistica

$$\frac{\overline{X}-\overline{Y}}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$$

Si accetta H_0 se

$$\frac{|\overline{X}-\overline{Y}|}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} < Z_{1-\frac{\alpha}{2}}$$

Abbiamo

$$\frac{|\overline{X}-\overline{Y}|}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} = \frac{|\overline{\mu}_1-\overline{\mu}_2|}{\text{"} \quad \text{"}} = \frac{11.15-10.20}{\sqrt{\frac{1.32}{9}+\frac{2.15}{16}}} = 1.792$$

D'altra parte

con $\alpha = 0.1$ $1-\frac{\alpha}{2} = 0.95$ e $Z_{0.95} = 1.65$

Quindi al livello $\alpha = 10\%$ H_0 viene respinta

con $\alpha = 0.05$ $1-\frac{\alpha}{2} = 0.975$ e $Z_{0.975} = 1.96$

Quindi al livello $\alpha = 5\%$ H_0 viene accettata

2 Il p-value è dato dalla relazione $Z_{1-\frac{p}{2}} = 1.79(2)$
da cui $1-\frac{p}{2} = \mathrm{Erf}(1.79) \Rightarrow p = 2\,(1-\mathrm{Erf}(1.79)) \simeq 8\%$.

Esame di *Calcolo delle probabilità e statistica* (per studenti di Informatica)
corso A

Università degli studi di Bari Aldo Moro

Docente: *Stefano Rossi*

09-09-2020

Esercizio 1. Si hanno due mazzi di chiavi. Il primo contiene n chiavi, il secondo $2n$ chiavi. Ciascun mazzo contiene esattamente una chiave che apre una porta, mentre le rimanenti chiavi non permettono di aprirla. Si sceglie a caso uno dei due mazzi e poi si estrae a caso una chiave dal mazzo prescelto. Si prova quindi ad aprire la porta. Se non si apre, si mette via la chiave che non funziona e se ne sceglie un'altra tra le restanti dello stesso mazzo e si riprova ad aprire la porta. Si continua finché non si trova la chiave che apre la porta.

(1) Calcolare la probabilità che sia necessario più di un tentativo per aprire la porta.
(2) Calcolare la probabilità che la chiave provenga dal primo mazzo sapendo che sono stati necessari due tentativi per aprire la porta (la seconda chiave scelta ha aperto la porta).
(3) Calcolare il limite per $n \to \infty$ della probabilità determinata al punto precedente.

Esercizio 2. Verificare che per ogni valore del parametro $\alpha > 0$ la funzione $f(x) = \frac{\alpha}{2} e^{-\alpha|x|}$, $x \in \mathbb{R}$, è la densità di probabilità di una certa variabile aleatoria X.

(1) Determinare lo stimatore di massima verosimiglianza di α corrispondente a un campione $(X_1, X_2, \dots, X_n)$ di rango n.
(2) Dopo aver verificato che $|X|$ è una variabile di tipo esponenziale di parametro α, calcolare il limite in probabilità della successione $\frac{|X_1|+|X_2|+\ldots+|X_n|}{n}$.

Esercizio 3. Due macchinari diversi dovrebbero produrre viti dello stesso diametro. La media $\bar{x}_1$ di un campione di $n_1 = 20$ viti prodotte dal primo macchinario è pari a $4.2\,mm$, mentre la media $\bar{x}_2$ di un campione di $n_2 = 25$ viti prodotte dal secondo è pari a $3.9\,mm$.

(1) Condurre un test di verifica delle ipotesi per stabilire se in media i due macchinari producono viti dello stesso diametro con un livello di significatività del 10% e del 5%, sapendo che la varianza delle due popolazioni di viti è uguale e pari a $\sigma^2 = 0.25\,mm^2$
(2) Calcolare il p-value del test di sopra.

(1) $M_1 =$ "Viene preso il primo mazzo"

$M_2 =$ "Viene preso il secondo mezzo"

$p(M_1) = p(M_2) = \frac{1}{2}$

1) $E =$ "È necessario più di 1 tentativo"

$p(E) = 1 - p(E^c)$

$E^c =$ "È necessario solo un tentativo"

$p(E^c) = p(E^c|M_1) \cdot p(M_1) + p(E^c|M_2) \cdot p(M_2)$ formula delle probabilità totale

$p(E^c|M_1) = \frac{1}{n}$

$p(E^c|M_2) = \frac{1}{2n}$

perciò $p(E^c) = \frac{1}{2n} + \frac{1}{4n} = \frac{2+1}{4n} = \frac{3}{4n}$

$p(E) = 1 - \frac{3}{4n}$

2) $F =$ "Sono necessari due tentativi"

$p(M_1|F) = \frac{p(F|M_1) \cdot p(M_1)}{p(F|M_2) \cdot p(M_2) + p(F|M_1) \cdot p(M_1)}$ formula di Bayes

$p(F|M_1) = \frac{n-1}{n} \cdot \frac{1}{n-1} = \frac{1}{n}$; $p(F|M_2) = \frac{2n-1}{2n} \frac{1}{2n-1} = \frac{1}{2n}$

perciò

$$p(F|M_1) = \frac{\frac{1}{n}\cdot\cancel{\frac{1}{2}}}{\frac{1}{n}\cdot\cancel{\frac{1}{2}} + \frac{1}{2n}\cdot\cancel{\frac{1}{2}}} = \frac{\frac{1}{n}}{\frac{2+1}{2n}} = \frac{1}{\cancel{n}}\cdot\frac{2\cancel{n}}{3} = \frac{2}{3}$$

non dipende da n !

3) $\lim_{n\to\infty} p(F|M_1) = \frac{2}{3}$

(2) $f(x) = \frac{\alpha}{2} e^{-\alpha|x|}$

$f(x) \geq 0 \quad \forall x \in \mathbb{R}$. Inoltre

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{\infty} \frac{\alpha}{2} e^{-\alpha|x|} dx = \cancel{2}\int_{0}^{\infty} \frac{\alpha}{\cancel{2}} e^{-\alpha x} dx$$

(perché f è una funzione pari)

$$= \alpha \left.\frac{e^{-\alpha x}}{-\alpha}\right|_0^{\infty} = -(0-1) = 1$$

(1) $f(x_1, x_2, \dots, x_n|\alpha) = \frac{\alpha}{2} e^{-\alpha|x_1|} \frac{\alpha}{2} e^{-\alpha|x_2|} \cdots \frac{\alpha}{2} e^{-\alpha|x_n|}$

$$= \left(\frac{\alpha}{2}\right)^n e^{-\alpha(|x_1|+|x_2|+\dots+|x_n|)}$$

$$\log f(x_1, x_2, \dots, x_n | \alpha) = n \ln\left(\frac{\alpha}{2}\right) - \alpha\left(|x_1| + |x_2| + \dots + |x_n|\right)$$

$$\frac{\partial}{\partial \alpha} \log f(x_1, \dots, x_n) = \frac{n}{\frac{\alpha}{2}} \cdot \frac{1}{2} - \left(|x_1| + |x_2| + \dots + |x_n|\right)$$

$$= \frac{2n}{\alpha} \cdot \frac{1}{2} - \sum_{i=1}^{n} |x_i| \overset{!}{=} 0$$

$$\alpha = \frac{n}{\sum_{i=1}^{n} |x_i|} = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} |x_i|}$$

$$\hat{\alpha} := \frac{1}{\frac{1}{n}\sum_{i=1}^{n} |X_i|}$$

è lo stimatore di massima verosimiglianza.

② Occorre calcolare la distribuzione di $|X|$.

Ora $P(|X| \leq t) = 0$ se $t < 0$.

se $t > 0$, invece, abbiamo

$$P(|X| \leq t) = P(-t \leq X \leq t) = \int_{-t}^{t} \frac{\alpha}{2} e^{-\alpha|x|}\, dx =$$

$\frac{\alpha}{2} e^{-\alpha|x_n|}$

$$= \int_{0}^{t} \alpha e^{-\alpha x}\, dx$$

Ma allora la densità g di $|X|$ è la fun

$$g(t) = \begin{cases} 0 & \text{se } t < 0 \\ \alpha e^{-\alpha t} & \text{se } t > 0 \end{cases}$$

Per la legge debole dei grandi numeri

$$\frac{|X_1| + |X_2| + \dots + |X_n|}{n}$$

converge debolmente a $E[|X|] = \frac{1}{\alpha}$

(3) $\overline{X}_1 = 4.2$ mm $\quad \overline{X}_2 = 3.9 \quad \sigma_1^2 = \sigma_2^2 = \sigma^2 = 0.25 \text{ mm}^2$

$n_1 = 20 \quad n_2 = 25 \quad \sigma = 0.5$ mm

1) $H_0 : \mu_1 = \mu_2$ con $\alpha = 0.1$ e $\alpha = 0.05$

Occorre considerare la statistica

$$\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}}$$

Rifiuto H_0 se $\left| \frac{\overline{X}_1 - \overline{X}_2}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right| > z_{1-\frac{\alpha}{2}}$

Ora $\left| \frac{X_1 - X_2}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right| = \frac{4.2 - 3.9}{0.5 \sqrt{\frac{1}{20} + \frac{1}{25}}} = 2$

Con $\alpha = 0.1 \quad 1 - \frac{\alpha}{2} = 0.95$ e $z_{0.95} = 1.65$

dato che $2 > 1.65$, rifiuto H_0

dunque rifiuto H_0 anche con $\alpha = 5\%$

2) Il p-value è determinato dalla condizione

$$z_{1-\frac{p}{2}} = 2$$

cioè

$$1 - \frac{p}{2} = Erf(2), \qquad p = 2(1 - Erf(2))$$

$$= 2(1 - 0.97725) = 0{,}0455$$

$$= 4.5\%$$

Esame di *Calcolo delle probabilità e statistica* (per studenti di Informatica) corso A

Università degli studi di Bari Aldo Moro
Docente: Stefano Rossi

28-09-2020

Esercizio 1. Si ha a disposizione un mazzo di $n \geq 1$ chiavi, che ne contiene solo una in grado di aprire una porta. Si conduce il seguente esperimento. Si sceglie una chiave a caso e si prova a vedere se apre la porta. Se non la apre, la si rimette nel mazzo e se ne estrae un'altra finché non si trova la chiave che apre la porta. Si lanciano quindi k monete equilibrate (le cui facce sono testa e croce), dove $k \geq 1$, è il numero di tentativi necessari ad aprire la porta.

(1) Calcolare la probabilità di ottenere almeno una testa.
(2) Sapendo di aver ottenuto almeno una testa, calcolare la probabilità che sia stato necessario solo un tentativo, verificando che tale probabilità è pari a $\frac{n+1}{2n^2}$.
(3) Calcolare il limite per $n \to \infty$ della probabilità determinata al punto precedente.

(Suggerimento: per il punto 1) è utile ricordare che $\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$.)

Esercizio 2. Sia $Y = 2X - 3$ una variabile aleatoria, dove X è una variabile di Poisson di parametro incognito $\lambda > 0$.

(1) Determinare l'aspettazione e la varianza di Y.
(2) Determinare lo stima di massima verosimiglianza per il parametro λ corrispondente alle 5 osservazioni di Y date da $y_1 = y_2 = -3$, $y_3 = 3$ e $y_4 = y_5 = 5$.

Esercizio 3. Due macchinari diversi dovrebbero produrre bottiglie che in media hanno lo stesso contenuto di vino. La media $\bar{x}_1$ di un campione di $n_1 = 25$ bottiglie prodotte dal primo macchinario è pari a $748cl$, mentre la media $\bar{x}_2$ di un campione di $n_2 = 25$ bottiglie prodotte dal secondo è pari a $751cl$. E' noto che la varianza del contenuto di tutte le bottiglie prodotte dal primo macchinario è $\sigma_1^2 = 30cl^2$, mentre per le bottiglie prodotte dal secondo macchinario si ha $\sigma_2^2 = 36cl^2$.

(1) Costruire un intervallo (bilatero) di confidenza per la differenza delle medie (reali) $\mu_1 - \mu_2$ con un livello di fiduca del 90% e del 95%.
(2) Condurre un test di verifica delle ipotesi per accertare che le due medie al punto di sopra siano uguali a un livello di significatività del 10% e del 5%.

Esercizio 1

H_K = " Sono necessari K tentativi per aprire la porta"

$$p(H_K) = q^{k-1} p$$

con $q = 1 - p$ e $p = \frac{1}{n}$ (n è il numero di chiavi)

1) E = " Si ottiene almeno una testa"

$$p(E) = 1 - p(E^C)$$

E^C = " Non si ottiene nemmeno una testa"

Ora

$$p(E^C) = \sum_{k=1}^{\infty} p(E^C | H_K) \cdot p(H_K)$$ (probabilità totale)

$$p(E^C | H_K) = \frac{1}{2^k}$$

per cui

$$p(E^C) = \sum_{k=1}^{\infty} \frac{1}{2^k} q^{k-1} \cdot p = \frac{p}{2} \sum_{k=1}^{\infty} \left(\frac{q}{2}\right)^{k-1} =$$

$$\frac{p}{2} \sum_{k=0}^{\infty} \left(\frac{q}{2}\right)^k = \frac{p}{2} \frac{1}{1-\frac{q}{2}} = \frac{p}{\cancel{2}} \cdot \frac{\cancel{2}}{2-q} =$$

$$= \frac{p}{2-(1-p)} = \frac{p}{1+p}$$

$$p(E) = 1 - \frac{p}{1+p} = \frac{1+\cancel{p}-\cancel{p}}{1+p} = \frac{1}{1+\frac{1}{n}} = \frac{n}{n+1}$$

2) $p(H_1 | E) = \frac{p(E|H_1) \cdot p(H_1)}{p(E)}$

$p(E | H_1) = \frac{1}{2}$

$p(H_1) = \frac{1}{n}$

perciò $p(H_1|E) = \frac{1}{2n} \frac{n+1}{n} = \frac{n+1}{2n^2}$

3) $\lim_{n \to \infty} p(H_1|E) = \lim_{n \to \infty} \frac{n+1}{2n^2} = 0$ (il denominatore è un polinomio di grado 2, mentre il numeratore è un polinomio di grado 1)

Esercizio 2

$Y = 2X - 3$ con X variabile di Poisson di parametro λ

1) $E[Y] = E[2X-3] = 2E[X] - 3 = 2\lambda - 3$

$Var[Y] = Var[2X-3] = Var[2X] = 4\,Var[X] = 4\lambda$

2) $Y_1 = Y_2 = -3 \qquad Y_3 = 3 \qquad Y_4 = Y_5 = 5$

$p(X=k) = \frac{e^{-\lambda}\lambda^k}{k!}$

$y = -3 \iff 2X - 3 = -3 \iff X = 0$ e $P(X=0) = e^{-\lambda}$

$y = 3 \iff 2X - 3 = 3 \iff X = 3$ e $P(X=3) = \frac{e^{-\lambda}\lambda^3}{3!}$

$y=5 \iff 2x-3=5; \iff x=4$ e $P(x=4)=\dfrac{e^{-\lambda}\lambda^4}{4!}$

la fun di verosimiglianza è allora

$$g(\lambda)=\underbrace{\left(e^{-\lambda}\right)^2}_{\text{è stato ottenuto 2 volte}} e^{-\lambda}\frac{\lambda^3}{3!}\underbrace{\left(\frac{e^{-\lambda}\lambda^4}{4!}\right)^2}_{\text{è stato ottenuto 2 volte}}$$

$$= e^{-2\lambda}\, e^{-\lambda}\,\frac{\lambda^3}{3!}\,\frac{e^{-2\lambda}\lambda^8}{(4!)^2}.$$

$$=\frac{e^{-5\lambda}\lambda^{11}}{3!\,(4!)^2}$$

Ora $h(\lambda):=\ln g(\lambda)= -5\lambda + 11\ln\lambda$ (a meno di costanti additive)

$h'(y)=-5+\dfrac{11}{\lambda}\overset{!}{=}0$; $-5\lambda=-11$; $\lambda=\dfrac{11}{5}=2.2$

$\lambda=2.2$ è la stima di massima verosimiglianza

(3) $\overline{X}_1 = 748\ cl$ $\quad n_1=25 \quad \sigma_1^2=30\ cl^2$

$\overline{X}_2 = 751\ cl$ $\quad n_2=25 \quad \sigma_2^2=36\ cl^2$

$\overline{x}_1-\overline{x}_2=-3$

Dato che le varianze sono note, l'intervallo di confidenza richiesto è della forma

$$\left[\overline{X}_1-\overline{X}_2 - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\ ;\ \overline{X}_1-\overline{X}_2 + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right]$$

Ora con $1-\alpha=0.9$ $\quad \alpha=0.1 \Rightarrow 1-\frac{\alpha}{2}=0.95$ e

$z_{0.95}=1.65$ $\quad z_{1-\frac{\alpha}{2}}\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}=1.65\sqrt{\dfrac{30+36}{25}}\simeq 2.681$

E l'intervallo è allora $[-5.681; -0.319]$

con $1-\alpha = 0{,}95$; $\frac{\alpha}{2} = 0.025$ e $1-\frac{\alpha}{2} = 0.975$

e $z_{0.975} = 1.96$

e allora $z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = 1.96 \sqrt{\frac{66}{25}} \cong 3.185$

e l'intervallo di fiducia corrispondente

$[-6{,}185 ; 0.185]$

(2) Senza effettuare ulteriori calcoli, $H_0: \mu_1 = \mu_2$ viene rifiutata con $\alpha = 0.1$ (perché 0 non è nel corrispondente intervallo di fiducia), ma accettata con $\alpha = 0.05$ (perché 0 appartiene al corrispondente intervallo di fiducia)

Esame di *Calcolo delle probabilità e statistica* (per studenti di Informatica) corso A

Università degli studi di Bari Aldo Moro

20-11-2020

Esercizio 1. Si deve stampare un documento di 3 pagine. Nel processo di stampa possono apparire dei refusi tipografici distribuiti su ciascuna pagina come una distribuzione di Poisson di parametri $\lambda_1 = 2$, $\lambda_2 = 1$ e $\lambda_3 = 3$ per la prima, la seconda, e la terza pagina, rispettivamente. Assumendo che la presenza di un refuso su una pagina sia indipendente dalla presenza di refusi sulle altre pagine, calcolare la probabilità che nel documento non ci sia complessivamente nemmeno un refuso e la probabilità che ce ne sia esattamente uno.
Si sceglie una pagina a caso.

(1) Calcolare la probabilità che contenga almeno un refuso.
(2) Sapendo che contiene esattamente due refusi, calcolare la probabilità che sia la seconda pagina.

Esercizio 2. Sia $\theta > 0$ un parametro assegnato. Verficare che la funzione $f(x;\theta) = \frac{2\theta^2}{x^3}$ per $x \geq \theta$ e 0 altrimenti è la densità di probabilità di una certa variabile aleatoria X.

(1) Calcolare l'aspettazione di X, verificando che si ha $E[X] = 2\theta$.
(2) Dato un campione $(X_1, X_2, \ldots, X_n)$ di rango n di variabili indipendenti e identicamente distribuite come X, usare il punto precedente per esibire uno stimatore corretto del parametro θ.
(3) Determinare lo stimatore di massima verosimiglianza di θ.

Esercizio 3.
Si conduce un test di verifica delle ipotesi sull'affermazione "la spesa media per le vacanze degli italiani è inferiore a 800 euro a persona". A tal scopo, si intervista un campione di 100 italiani e si trova che la spesa media per le vacanze di questo campione è stata di 810 euro a persona con una varianza campionaria pari a $S^2 = 1600$ euro.

(1) Dire se bisogna condurre un test unilaterale o bilaterale, specificando di quale test si tratta.
(2) Effettuare il test con un livello di significatività del 5% e dell' 1%.

Esercizio 1

Indichiamo con X_i, $i=1,2,3$, le v.a. che contano il numero di refusi che appaiono nella pagina i.
X_1 è di Poisson di parametro $\lambda_1 = 2$
X_2 è di Poisson di parametro $\lambda_2 = 1$
X_3 è di Poisson di parametro $\lambda_3 = 3$

Indichiamo con $X := X_1 + X_2 + X_3$ la variabile aleatoria che conta il numero complessivo di refusi.
Dato che X_1, X_2, X_3 sono indipendenti, X è ancora una variabile di Poisson, di parametro $\lambda = \lambda_1 + \lambda_2 + \lambda_3 = 2+1+3 = 6$.

Ora $p(X=k) = e^{-\lambda} \frac{\lambda^k}{k!}$ per ogni $k \in \mathbb{N}$, per cui

$$p(X=0) = e^{-6} \frac{6^0}{0!} = \frac{1}{e^6} \simeq 2.48 \cdot 10^{-3} = 0.00248$$

$$p(X=1) = e^{-6} \, 6^1 = \frac{6}{e^6} = 0.015$$

1. $H_i =$ "Scelgo la pagina i" $\quad i = 1,2,3$

Si suggerisca di assumere gli eventi H_i equiprobabili, cioè
$p(H_1) = p(H_2) = p(H_3) = \frac{1}{3}$

$E =$ "La pagina scelta contiene almeno un refuso"

$p(E) = 1 - p(E^c)$ con

$E^c =$ "La pagina scelta non contiene nemmeno un refuso"

$$p(E^c) = p(H_1)\, p(E^c|H_1) + p(H_2) \cdot p(E^c|H_2) + p(H_3) \cdot p(E^c|H_3)$$
$$= \frac{1}{3} \left[p(X_1=0) + p(X_2=0) + p(X_3=0) \right] = \frac{1}{3} \left(e^{-\lambda_1} + e^{-\lambda_2} + e^{-\lambda_3} \right)$$

$$= \frac{1}{3}(e^{-2} + e^{-1} + e^{-3}) = \frac{1}{3}\left(\frac{1}{e^2} + \frac{1}{e} + \frac{1}{e^3}\right) = 0.184$$

Quindi $p(E) = 0.816$

2. $F =$ "La pagina contiene 2 errori)

$$p(H_2|F) = \frac{p(F|H_2)\cdot \cancel{p(H_2)}}{p(F|H_1)\cdot \cancel{p(H_1)} + p(F|H_2)\cdot \cancel{p(H_2)} + p(F|H_3)\cdot \cancel{p(H_3)}} =$$

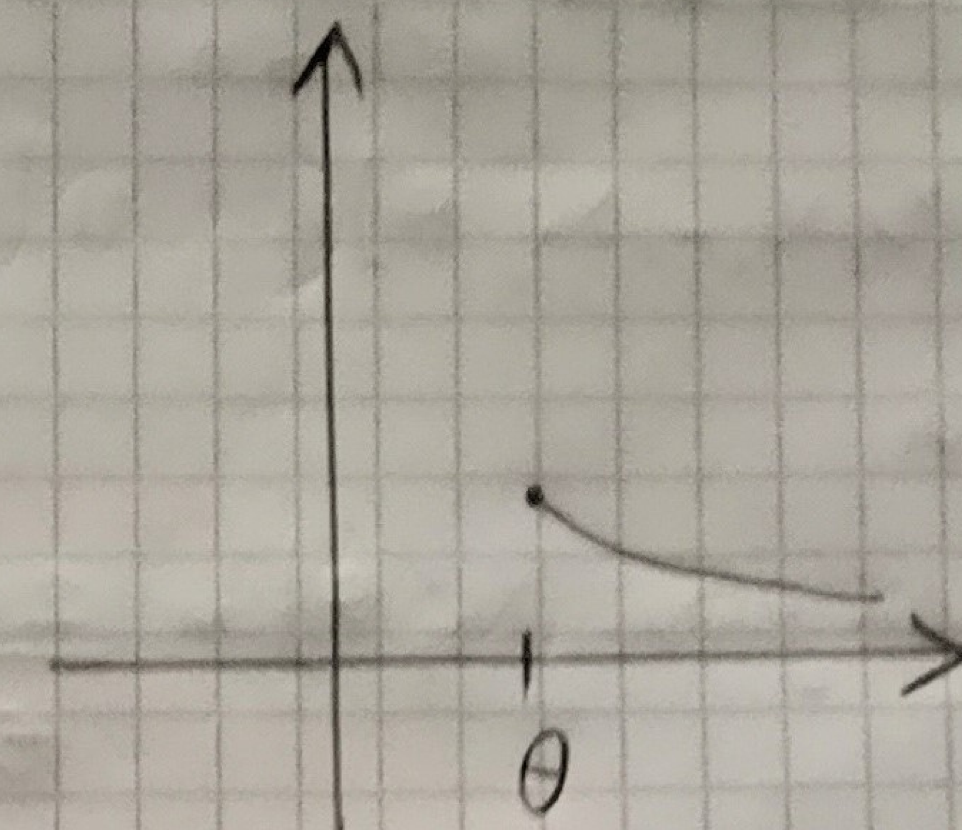
$$\frac{p(F|H_2)}{p(F|H_1) + p(F|H_2) + p(F|H_3)} =$$

$$\frac{P(X_2 = 2)}{P(X_1 = 2) + P(X_2 = 2) + P(X_3 = 2)} = \frac{e^{-1}\cdot\frac{1^2}{2!}}{e^{-2}\cdot\frac{2^2}{\cancel{2!}} + e^{-1}\cdot\frac{2^1}{2!} + e^{-3}\,\frac{3^2}{\cancel{2!}}}$$

$$= \frac{\frac{1}{e}}{\frac{4}{e^2} + \frac{1}{e} + \frac{9}{e^3}} \simeq 0.5$$

Esercizio 2

$$f(x;\theta) = \begin{cases} \frac{2\theta^2}{x^3} & x \geq \theta \\ 0 & x < \theta \end{cases}$$



Per prima cosa, osserviamo che $f(x,\theta) \geq 0$ $\forall x \in \mathbb{R}$

Inoltre $\int_{-\infty}^{\infty} f(x,\theta)\,dx = \int_{\theta}^{+\infty} \frac{2\theta^2}{x^3}\,dx = 2\theta^2 \int_{\theta}^{\infty} \frac{dx}{x^3} = 2\theta^2 \lim_{M\to+\infty} \int_{\theta}^{M} x^{-3}\,dx =$

Ora

$$= 2\theta^2 \left[\frac{x^{-3+1}}{-3+1} \right]_{\theta}^{M} = 2\theta^2 \left[-\frac{1}{2x^2} \Big|_{\theta}^{M} \right] = 2\theta^2 \left[-\frac{1}{2M^2} + \frac{1}{2\theta^2} \right]$$

$$\xrightarrow[M\to\infty]{} \frac{2\theta^2}{2\theta^2} = 1$$

Quindi la funz. $f(x;\theta)$ è normalizzata per ogni valore di θ.

1. $E[X] := \int_{\theta}^{\infty} x \frac{2\theta^2}{x^3}\,dx = 2\theta^2 \lim_{M\to+\infty} \int_{\theta}^{M} x^{-2}\,dx =$

$$= 2\theta^2 \lim_{M\to\infty} \left[\frac{x^{-2+1}}{-2+1} \Big|_{\theta}^{M} \right] = 2\theta^2 \lim_{M\to+\infty} \left[-\frac{1}{x} \Big|_{\theta}^{M} \right] =$$

$$2\theta^2 \lim_{M\to+\infty} \left(-\frac{1}{M} + \frac{1}{\theta} \right) = \frac{2\theta^2}{\theta} = 2\theta.$$

2. Dato che $\theta = E\left[\frac{X}{2}\right]$, segue facilmente che

$$T_n := \frac{1}{2} \frac{X_1 + X_2 + \dots + X_n}{n}$$ è uno stimatore corretto di θ

$\left(\text{Infatti, } E[T_n] = \frac{1}{2n} \sum_{i=1}^{n} E[X_i] = \frac{1}{2n} n \cdot 2\theta = \theta \right)$

3. La densità congiunta di $(X_1, X_2, \dots, X_n)$ è data dal prodotto

$$g(x_1, x_2, \dots, x_n; \theta) = 2^n \theta^{2n} x_1^{-3n} x_2^{-3n} \dots x_n^{-3n}$$

per $x_i \geq \theta$ $i = 1, \dots, n$. Quindi $\theta \leq x_i$ $\forall i = 1, \dots, n$

cioè $0 < \theta \le \min_{1 \le i \le n} x_i$

$$\ln g(x_1, x_2, \dots, x_n; \theta) = n \ln 2 + 2n \ln \theta - 3n \ln x_1 \dots - 3n \ln x_n$$

per cui

$$\frac{\partial}{\partial \theta} \ln g(x_1, \dots, x_n; \theta) = \frac{2n}{\theta} > 0 \quad \forall \theta$$

ne segue che il massimo della f.ne g (come funzione di θ) è assunto nell'estremo destro dell'intervallo dove varia θ $\left(\theta \in \left(0, \min_{1 \le i \le n} x_i\right)\right)$, cioè $\theta = \min_{1 \le i \le n} x_i$

Lo stimatore di massima verosimiglianza sarà allora

$$\hat{\theta}_n = \min_{1 \le i \le n} \{X_i\}.$$

Esercizio 3

$H_0: \mu \leq \mu_0$ con $\mu_0 = 800$ (euro)

$n = 100$

$\overline{X} = 810$

$S^2 = 1600$

1) Si tratta di un test unilaterale di Student, perché la varianza non è nota.

La statistica da utilizzare è allora

$$T = \frac{\overline{X} - \mu_0}{\frac{S}{\sqrt{n}}} \quad \text{con } \overline{X} = \frac{X_1 + \dots + X_n}{n} \quad \text{(media campionaria)}$$

che è distribuita come una t di Student con $n-1 = 99$ gradi di libertà.

2) La regione di rifiuto è data da

$$T > t_{1-\alpha}(n-1)$$

Nel nostro caso

$$T = \frac{810 - 800}{\frac{40}{10}} = \frac{10}{4} = 2.5$$

- Ora con $\alpha = 5\%$ $1-\alpha = 0.95$ e $t_{0.95}(99) = 1.660$
per cui H_0 si rifiuta

- Con $\alpha = 1\%$ $1-\alpha = 0.99$ e $t_{0.99}(99) = 2.365$
per cui H_0 si rifiuta