STATISTICA A – K (63 ore)

Marco Riani

mriani@unipr.it

http://www.riani.it

Esercizio

Un partito politico ha commissionato un’indagine sull’orientamento della popolazione al prossimo referendum. Al partito interessa sapere se la percentuale dei votanti è la stessa nelle regioni chiave A e B. Nella regione A su 500 intervistati 300 hanno dichiarato che voteranno sì, nella regione B, su 600 intervistati, 340 hanno dichiarato che voteranno sì.

Si definisca l’ipotesi nulla e l’ipotesi alternativa e si dica quali conclusioni si ottengono assumendo $\alpha=0,05$. Si calcoli il p-value del test.

Soluzione

- H_0: $\pi_1 = \pi_2$
- H_1: $\pi_1 \neq \pi_2$ $\Rightarrow$ (la percentuale di coloro che voteranno sì è diversa nelle due regioni chiave)

Regione 1	Regione 2
$n_1 = 500$	$n_2 = 600$
$p_1 = 300/500 = 0{,}6$	$p_2 = 340/600 = 0{,}567$

Il test si basa su $DP = P_1 - P_2$

Utilizzando il Teorema centrale del limite

se è vera l'ipotesi nulla

$$Z(DP) = \frac{P_1 - P_2}{\sqrt{VAR(DP)}} \sim N(0,1)$$

$$Z(DP) = \frac{P_1 - P_2}{s(DP)} \sim N(0,1)$$

Soluzione

- H_0: $\pi_1=\pi_2$
- H_1: $\pi_1\neq\pi_2$ $\Rightarrow$ **(la percentuale di coloro che voteranno sì è diversa nelle due regioni chiave)**

	Regione 1	Regione 2
•	$n_1 = 500$	$n_2 = 600$
•	$p_1 = 0{,}6$	$p_2 = 0{,}567$

$$Z(DP) = \frac{P_1 - P_2}{\sqrt{VAR(DP)}} \sim N(0,1)$$

$$\mathrm{s(DP)} = \sqrt{p(1-p)\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}$$

$$\mathrm{s(DP)} = \sqrt{0{,}5818(1-0{,}5818)\left[\frac{1}{500}+\frac{1}{600}\right]} = 0{,}02987$$

$$p = \frac{p_1 n_1 + p_2 n_2}{n_1 + n_2} \quad \text{Stima di } \pi$$

$$\frac{0{,}6\times 500 + 0{,}567\times 600}{1100} = 0{,}5818$$

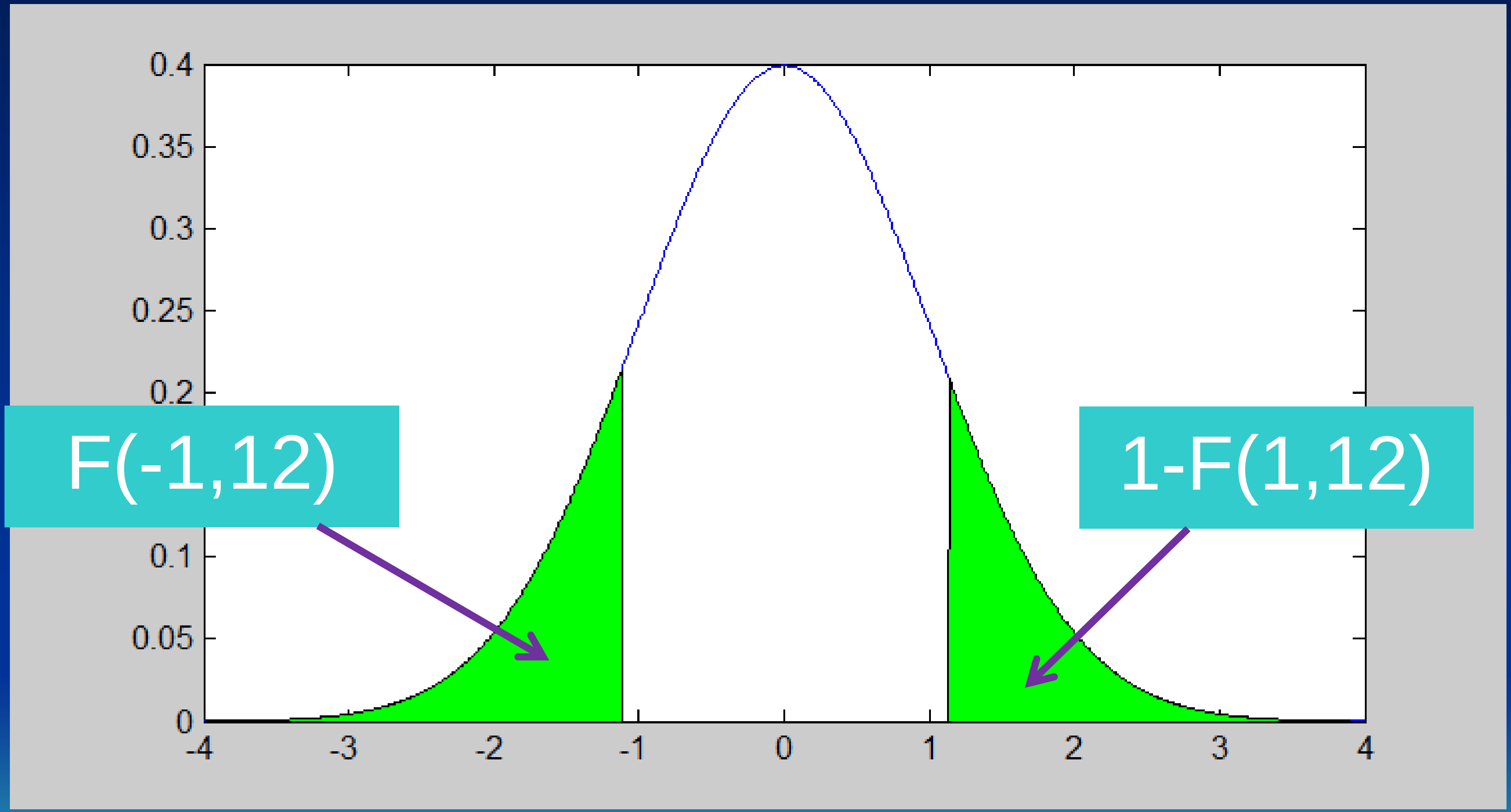
Calcolo del valore del test

$$Z(DP) = \frac{P_1 - P_2}{\sqrt{VAR(DP)}} \sim N(0,1)$$

$$Z(DP) = \frac{300/500 - 340/600}{0,02987} = 1,12$$

- **Assumendo α=0,05 la zona di accettazione è compresa nell'intervallo -1,96 1.96**
- **Il valore trovato cade nella zona di accettazione**

Calcolo del p-value di 1.12



-1,12 0 1,12

F(-1,12)+1-F(1,12)= 2(1-F(1,12))=0,2627

Esercizio

- Una moneta viene lanciata 80 volte, ottenendo 45 volte l’esito «testa».
- Al livello di significatività del 5% vi è sufficiente evidenza per ritenere che la moneta sia truccata?

Soluzione

H_0: π = 0,5 (la moneta non è truccata)

H_1: $\pi \neq$ 0,5 (la moneta è truccata)

Livello di significatività α = 0,05

Campione di 80 lanci (n=80):

p = 45/80=0,5625

$$Z(P) = \frac{P - \pi_0}{\sigma(P)} \sim N(0,1)$$

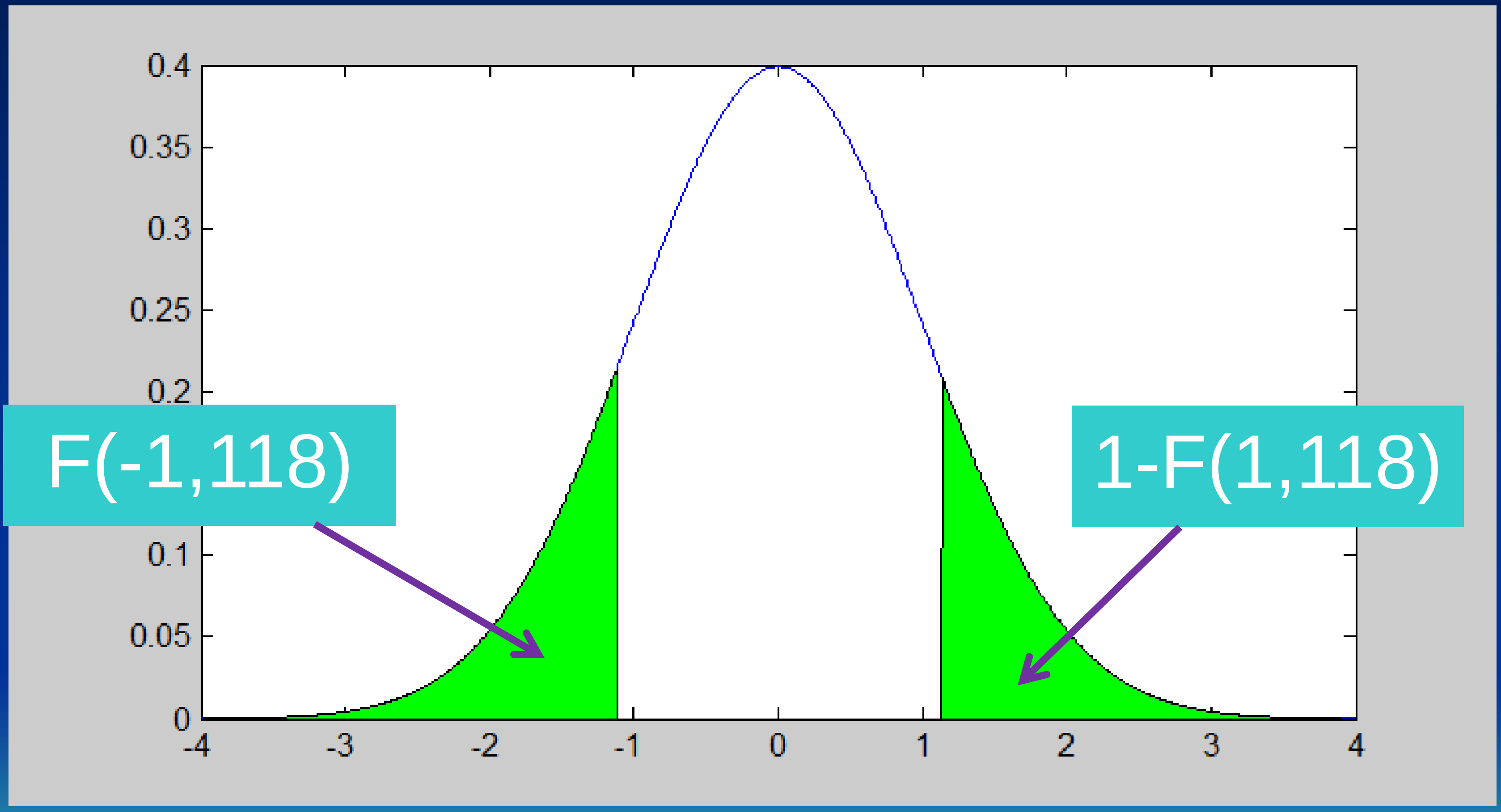
$$\sigma(p) = \sqrt{\frac{\pi_0(1-\pi_0)}{n}} = \sqrt{\frac{0,5 \cdot 0,5}{80}} = 0,0559$$

$$z(p) = \frac{0,5625 - 0,5}{0,0559} = 1,118$$

- Assumendo α=0,05 la zona di accettazione è compresa nell’intervallo -1,96 1.96
- Il valore trovato 1,118 cade nella zona di accettazione

P value del test?

Calcolo del p-value



-1,12 **0** **1,12**

F(-1,118)+1-F(1,118)= 2(1-F(1,118))=0,2336

Esercizio

- Nel processo di controllo del peso delle confezioni di un determinato prodotto l’azienda esamina un campione di 800 confezioni e trova che 15 di esse hanno un peso fuori norma.
- Si determini l’intervallo di confidenza al 97% della proporzione di pezzi fuori norma.
- Si testi, al livello di significatività dell'1%, l'ipotesi che la proporzione di pezzi fuori norma sia pari a 1,5%.
- Se la proporzione di pezzi fuori norma nell'universo fosse uguale a 1,5%, effettuando cinque estrazioni
  - si calcoli la probabilità di trovare esattamente due pezzi fuori norma;
  - si scriva l'espressione che consente di calcolare la probabilità di ottenere un numero di pezzi fuori norma compreso tra due e quattro (estremi compresi).

Intervallo di confidenza al 97%

- ***n* = 800 confezioni; 15 fuori norma**
- ***p* = 15/800 =** 0,01875 $\Rightarrow$ **1,875% (stima campionaria di π)**
- ***α=0,03 ➔ F(0,985)=2,17***

$$P\left\{p - z(\alpha/2)s(p) \le \pi \le p + z(\alpha/2)s(p)\right\} = 1-\alpha$$

$$s(p) = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0{,}01875(1-0{,}01875)}{800}} = 0{,}005$$

$$P\left\{0{,}00834 \le \pi \le 0{,}0292\right\} = 0{,}97$$

Test ipotesi sulla difettositià

H_0: $\pi = 0{,}015$
H_1: $\pi > 0{,}015$ **(difettosità superiore)**

Campione di 800 confezioni

p = 0,01875 **Fisso $\alpha = 0{,}01$**

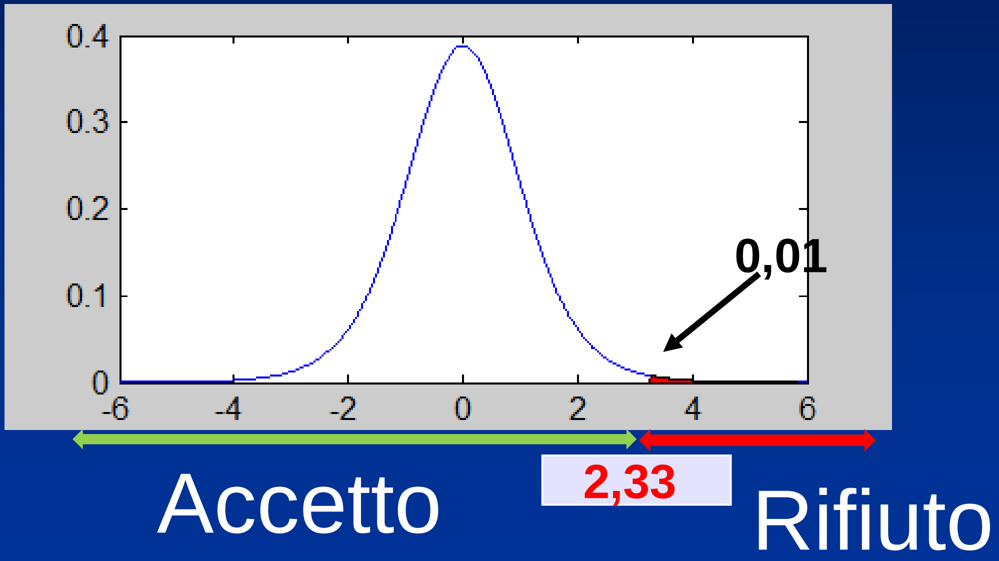
$$\sigma(P) = \sqrt{\frac{\pi_0(1-\pi_0)}{n}} = \sqrt{\frac{0{,}015 \cdot (1-0{,}015)}{800}} = 0{,}0039$$

$$Z(P) = \frac{P - \pi_0}{\sigma(P)} \sim N(0{,}1)$$

$$z(p) = \frac{0{,}01875 - 0{,}015}{0{,}0039} = 1{,}59$$

H_1: $\pi > 0,015$

α=0,01 F (2,33)=0,99



$$z(p) = 1,59$$

Con **α=0,01 i**l valore osservato del test cade nella zona di accettazione

Se la proporzione di pezzi fuori norma nell'universo fosse uguale a 1,5%, effettuando **cinque estrazioni**

– si calcoli la probabilità di trovare esattamente due pezzi fuori norma;

$$P(S=2) = \binom{5}{2} 0{,}015^2 (1-0{,}015)^3 = 0,00215$$

– si scriva l'espressione che consente di calcolare la probabilità di ottenere un numero di pezzi fuori norma compreso tra due e quattro (estremi compresi).

$$\sum_{i=2}^{4} Pr(S=i) = \sum_{i=2}^{4} \binom{5}{i} 0,015^i (1-0,015)^{5-i} = 0,002183$$

Esercizio

Un tipo di componente viene fornito in confezioni da 400 pezzi. Ne testiamo un campione di 16 per stimare la frazione di difettosi: vogliamo fare un test al livello di significatività α del 5% che ci permetta di rifiutare l'intera partita se vi è evidenza statistica che i pezzi difettosi (nella confezione) sono più del 15%

Quesiti

- Qual `e il parametro incognito su cui basare il test? Come vanno scelte ipotesi nulla e alternativa? Se nel campione si trovano 3 difettosi, cosa si decide? Quanti difettosi si possono accettare al massimo nel campione senza rifiutare la fornitura?
- Se una confezione ha il 25% di difettosi, con che probabilità questo test la rifiuta?

Soluzione

- Qual `e il parametro incognito su cui basare il test?
- π = ignota frazione di difettosità del lotto
- Come vanno scelte ipotesi nulla e alternativa?
- H_0: $\pi = \pi_0 = 0.15$
- H_1: $\pi > \pi_0 = 0.15$

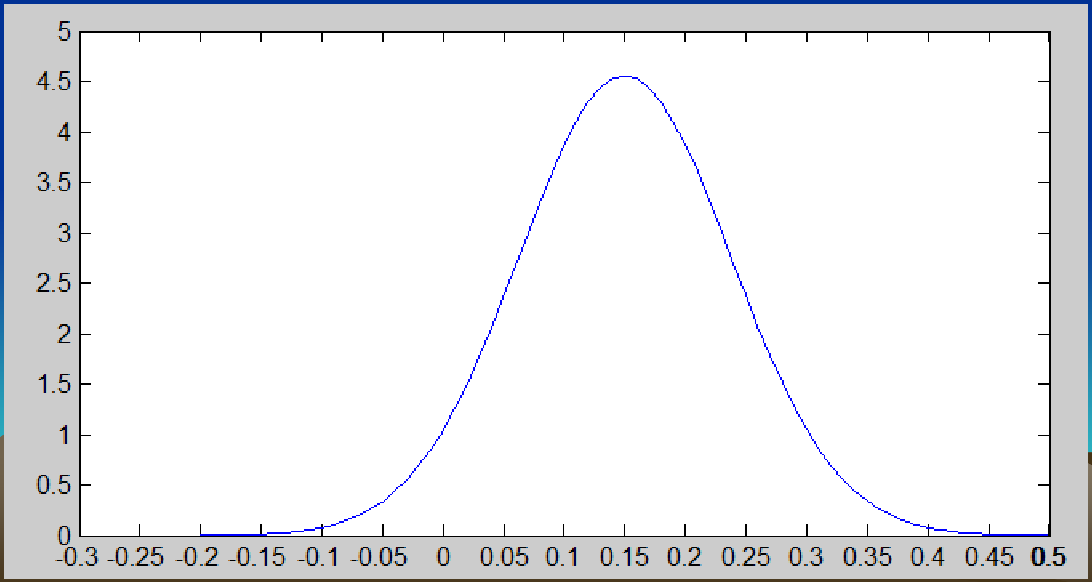
$H_0: \pi = \pi_0 = 0.15$
$H_1: \pi > \pi_0 = 0.15$

- Statistica test
- P = v.c. frequenza relativa campionaria
- Distribuzione di P (universi infiniti)
- P~N(π π(1- π)/n)
- Distribuzione di P (universi finiti)
- P~N(π   π(1- π)/n (N-n)/(N-1))

- In questo esempio N=400 (confezioni da 400 pezzi)

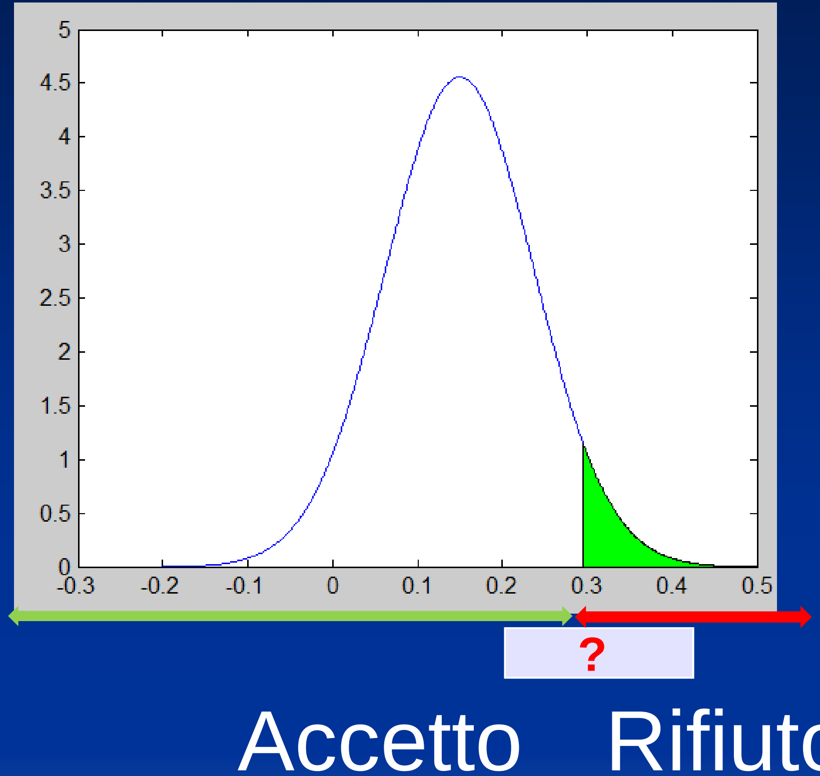
$H_0: \pi = \pi_0 = 0.15$
$H_1: \pi > \pi_0 = 0.15$

- Se è vera l'ipotesi nulla e testiamo un campione di 16 pezzi da un lotto di 400
- P~N(0.15 0.15(1- 0.15)/16 (400-16)/(400-1))
- P~N(0.15 0.0077)



Obiettivo: vogliamo fare un test al livello di significatività α del 5% che ci permetta di rifiutare l'intera partita se vi è evidenza statistica che i pezzi difettosi (nella confezione) sono più del 15%

Qual è la difettosità massima tollerabile nel campione se α=0.05



- Obiettivo: trovare il quantile $x_{0.95}$ nella distribuzione
- P~N(0.15 0.0077)

$$\frac{\bar{x}_\alpha - 0.15}{\sqrt{0.0077}} = 1{,}645$$

Il valore soglia $\bar{x}_\alpha$ è 0,294

Numero massimo di pezzi difettosi tollerabili nel campione = 0,294 ×16=[4.70]=4

Se nel campione si trovano 3 difettosi, cosa si decide?

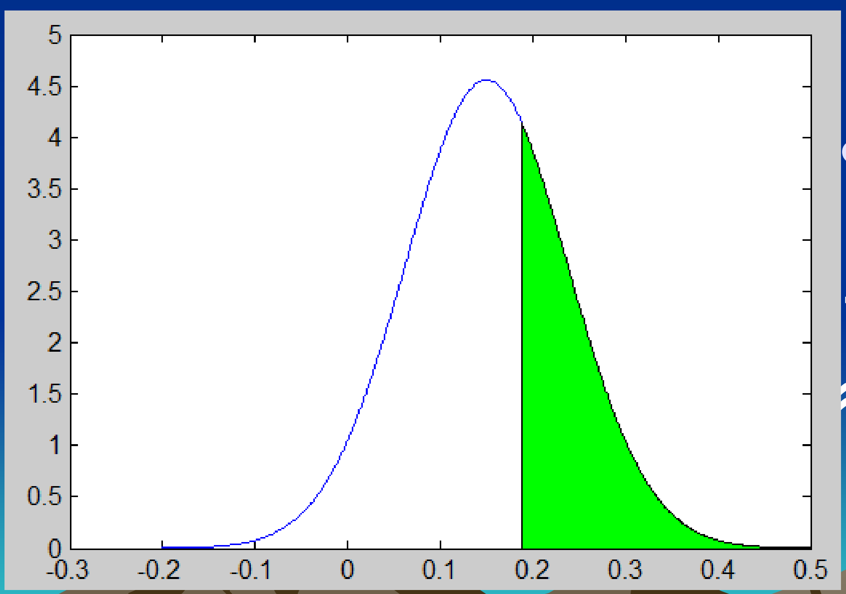
Numero massimo di pezzi difettosi tollerabili nel campione = 0,294 × 16=[4.70]=4

Se numero di difettosi nel campione =3 accetto l'ipotesi nulla

Qual è il p-value di 3 pezzi difettosi?

P-value di 3

- P-value di 3/16=0.1875
- Pr di trovare un valore superiore a 0.1875 in P~N(0.15 0.0077)



- P-value =

$1-F((0.1875-0.15)/0.0077^{0.5})$

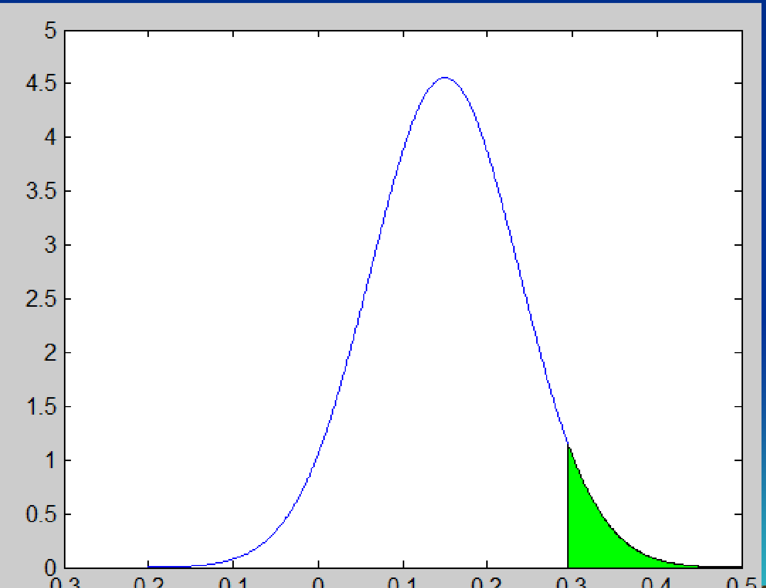
≈ 0.33

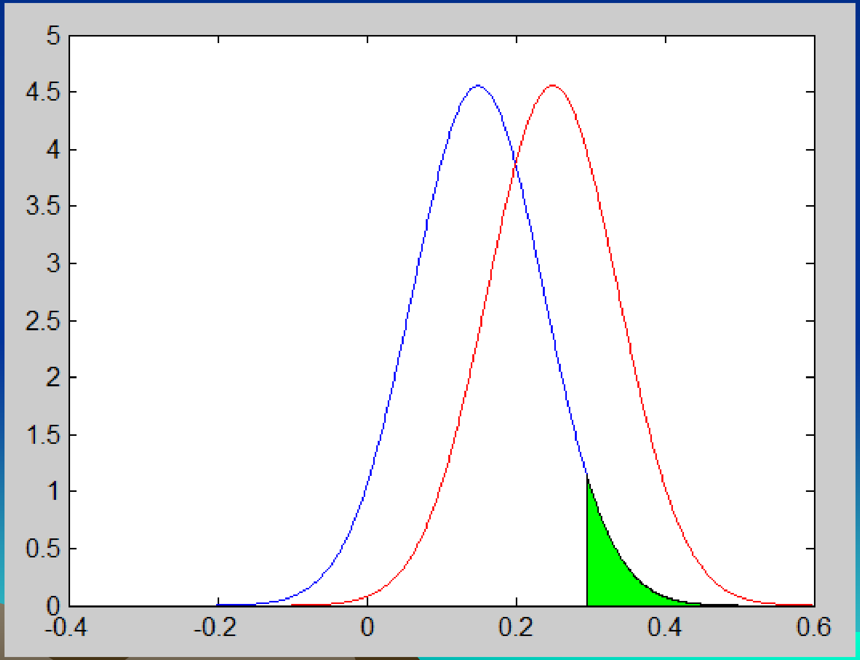
Se una confezione ha il 25% di difettosi, con che probabilità questo test la rifiuta?

Obiettivo: trovare la probabilità in una v.c.

P~N(0.25 0.15(1- 0.15)/16 (400-16)/(400-1))

di ottenere valori che cadono nella zona di rifiuto del test





0.294

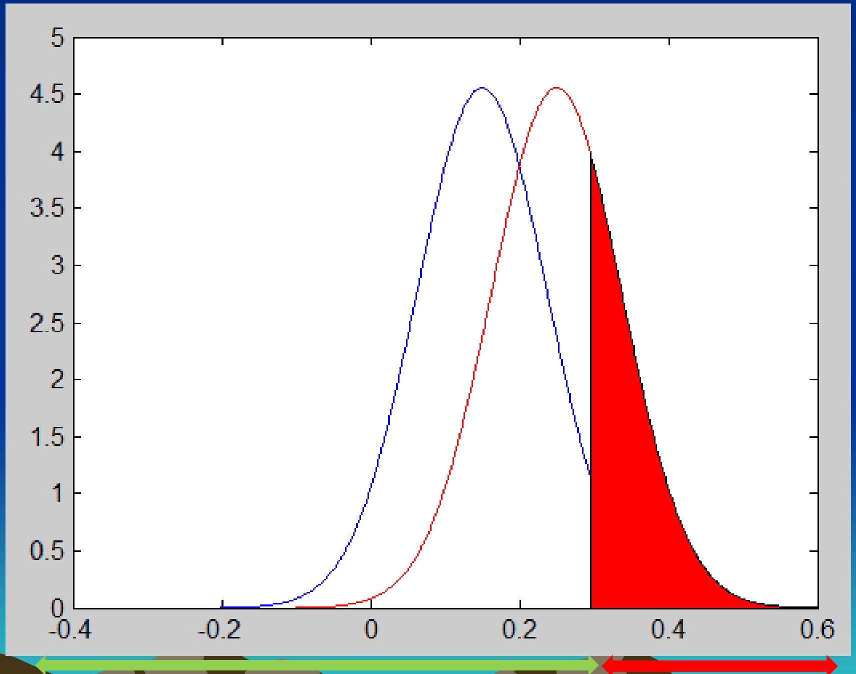
Accetto Rifiuto

Potenza del test quando π=0.25

Obiettivo: trovare la probabilità in una v.c.

$P \sim N(0.25 \quad 0.15(1- 0.15)/16 \ (400-16)/(400-1))$

di ottenere valori che cadono nella zona di rifiuto del test



0.294

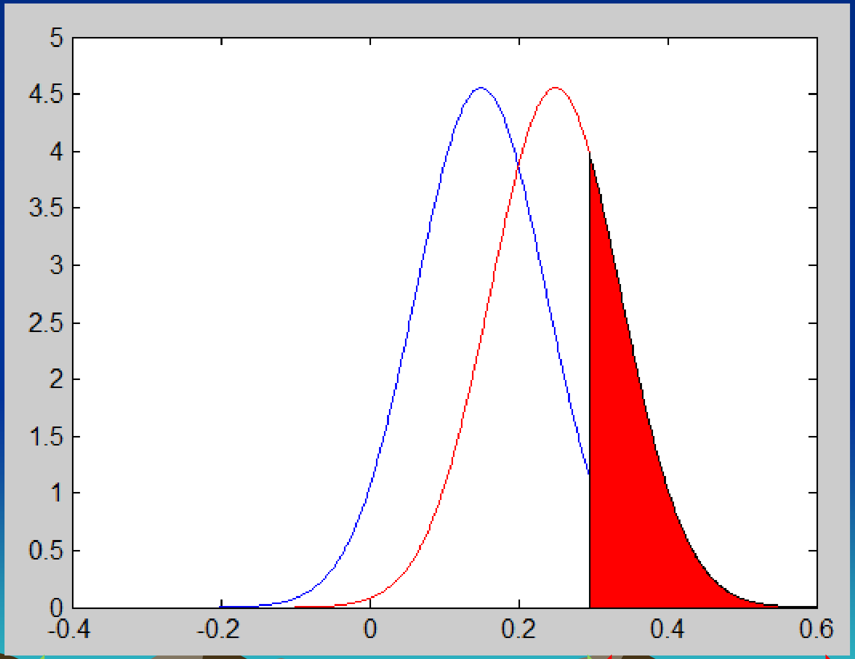
Accetto

Rifiuto

Potenza del test quando π=0.25

P~N(0.25 0.0077)

$1\text{-}F((0.294-0.25)/0.0077^{0.5}) \approx 0.308$



0.294

Accetto

Rifiuto

Esercizio

- Si consideri un dado a 20 facce tutte uguali
- Qual è il valore atteso?
- Quante volte è necessario lanciarlo affinché la probabilità di ottenere almeno un 20 sia maggiore o uguale a 0.5?
- Lanciandolo 20 volte, qual è il numero medio di 20 ottenuti?
- Pr di ottenere almeno una volta la faccia 20 in 20 lanci?

Soluzione

- Valori assunti da X
- i= 1, 2, …, 20
- P(X=i)=1/20 E(X)?

$$E(X) = \sum_{i=1}^{20} i \frac{1}{20} = \frac{1}{20} \frac{1+20}{2} 20 = 10{,}5$$

Quante volte è necessario lanciarlo affinché la probabilità di ottenere almeno un 20 sia maggiore o uguale a 0.5?

- Pr(almeno un 20 in n lanci) =
- 1-Pr(nessun 20 in n lanci)=
- $1-(19/20)^n=1-0.95^n$
- Vincolo
- $1-(19/20)^n>0.5$
- $n > \ln(0.5)/\ln(19/20)=\log_{0.95} 0.5=13.51\approx 14$

Lanciandolo 20 volte, qual e il numero medio di 20 ottenuti?

- Y_i v.c. associata all'ottenimento del numero 20
- Y_i ~ Bernoulliana π= 1/20
- $S_{20} = Y_1 + \ldots + Y_{20} \sim B(20, 1/20)$

- $E(S_{20}) = 20\,\pi = 20\,(1/20) = 1$

Pr di ottenere almeno una volta la faccia 20 in 20 lanci?

- Pr(almeno un 20 in 20 lanci) =
- 1-Pr(nessun 20 in 20 lanci)=
- 1-(1-1/20)^{20}=

$$1-\left(1-\frac{1}{20}\right)^{20}=0.641514$$

Nota: il valore di $(1-\frac{1}{20})^{20}=0.358486$ è vicino ad $e^{-1}=0.367879$.

Esercizio

- Nel gioco del lotto un numero ha una probabilità p di uscire ad ogni estrazione.
- Si scriva la densità della v.c. (X) che descrive il tempo di attesa dell’uscita del numero all’estrazione k-esima (v. casuale geometrica), k=1, 2, 3, ….
- Si dimostri che la somma delle probabilità è 1
- Si calcoli il valore atteso e la varianza di X
- Si calcoli l’espressione che definisce P(X>k)

Soluzione

- Dato che le estrazioni sono indipendenti la prob. che il numero esca alla k-esima estrazione è
- $P(X=k) = ?$ $k=1, 2, \ldots.,$
- $P(X=k) = (1-p)^{k-1}\,p$ $k=1, 2, \ldots.,$
- $P(X=k) = q^{k-1}\,p$ $k=1, 2, \ldots.,$

Si dimostri che la somma delle probabilità è 1

$$\sum_{k=1}^{\infty} P(X=k) = \sum_{k=1}^{\infty} (1-p)^{k-1} p$$

$$\sum_{k=1}^{\infty} (1-p)^{k-1} p = \sum_{k=0}^{\infty} (1-p)^{k} p$$

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

Valore atteso

q=1-p   P(X=k) = q^{k-1} p   k=1, 2,

$$E(X) = \sum_{k=1}^{\infty} kP(X=k) = p\sum_{k=0}^{\infty} k(1-p)^{k-1}$$

$$= p\sum_{k=0}^{\infty} kq^{k-1}$$

$$= p\sum_{k=0}^{\infty} \frac{dq^k}{dq}$$

$$= p\frac{d}{dq}\sum_{k=0}^{\infty} q^k$$

$$= p\frac{d}{dq}\frac{1}{1-q} =$$

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

Varianza= $E(X^2)-[E(X)]^2$

- Cominciamo a calcolare $E(X(X-1))$

$q=1-p \qquad P(X=k) = q^{k-1}\, p \qquad k=1, 2,$

$$E(X(X-1)) = p\sum_{k=1}^{\infty} k(k-1)(1-p)^{k-1} = p\sum_{k=0}^{\infty} k(k-1)q^{k-1} =$$

$$E(X(X-1)) = pq\sum_{k=0}^{\infty} k(k-1)q^{k-2}$$

$$= pq\sum_{k=0}^{\infty} \frac{d^2 q^k}{dq^2}$$

$$= pq\frac{d^2}{dq^2}\sum_{k=0}^{\infty} q^k =$$

$$= pq\frac{d^2}{dq^2}\frac{1}{1-q}$$

$$= pq\frac{2}{(1-q)^3} = \frac{2q}{p^2}$$

$$E(X^2) = \frac{2q}{p^2} + \frac{1}{p}$$

$$\mathrm{var}(X) = E(X^2) - [E(X)]^2 = \frac{q}{p^2}$$

Calcolare P(X>k)

- $P(X=k) = q^{k-1}\,p \quad k=1, 2, 3, \ldots$
- P(X>k) si può scrivere come

$$1-[P(X=1)+P(X=2)+P(X=3)+\ldots+P(X=k)]$$

$$=1-[p+qp+q^2p+\ldots+q^{k-1}p]=$$

$$=1-p\frac{1-q^k}{1-q}=q^k$$

Esercizio

- Dimostrare che nel gioco del lotto la probabilità che siano necessari i+j tentativi prima di ottenere il primo successo, dato che ci sono già stati i insuccessi consecutivi, è uguale alla probabilità non condizionata che almeno j tentativi siano necessari prima del primo successo.
- *Morale: il fatto di avere già osservato i insuccessi consecutivi non cambia la distribuzione del numero di tentativi necessari per ottenere il primo successo*

Soluzione

- X = numero di tentativi prima di ottenere il primo successo.
- p = prob di successo
- Dobbiamo dimostrare che
- $P(X>i+j \mid X>j) = P(X>i)$
- $P(X>i+j \mid X>j) = P(X>i+j \cap X>j) / P(X>j)$
- $= P(X>i+j) / P(X>j)$
- $= q^{i+j}/q^j=q^i=P(X>i)$

Esercizio

- Sia X una v.c. definita nell'intervallo [0 +∞)

$$f_X(x) = cxe^{-\frac{x^2}{2}}$$

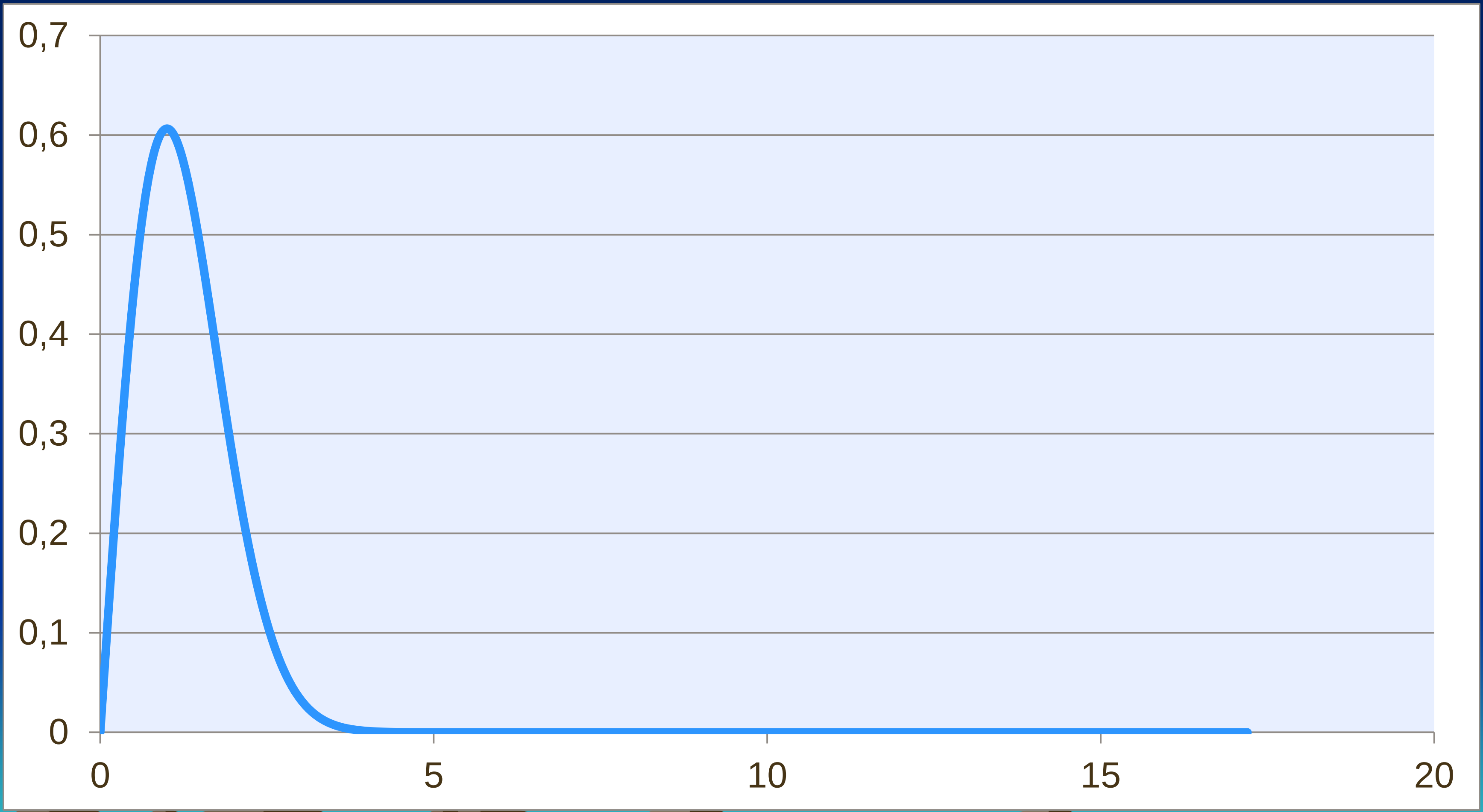
- Calcolare il valore di c affinché $f_X(x)$ sia effettivamente una densità
- Rappresentarla graficamente la funzione di densità
- Calcolare la funzione di ripartizione e rappresentarla graficamente
- Calcolare P(X>x)

Soluzione

$$1 = \int_0^{+\infty} Cxe^{-\frac{x^2}{2}}\ dx = [-Ce^{-\frac{x^2}{2}}]_0^{+\infty} = C$$

- Quindi C=1

Rappresentazione grafica della funzione di densità



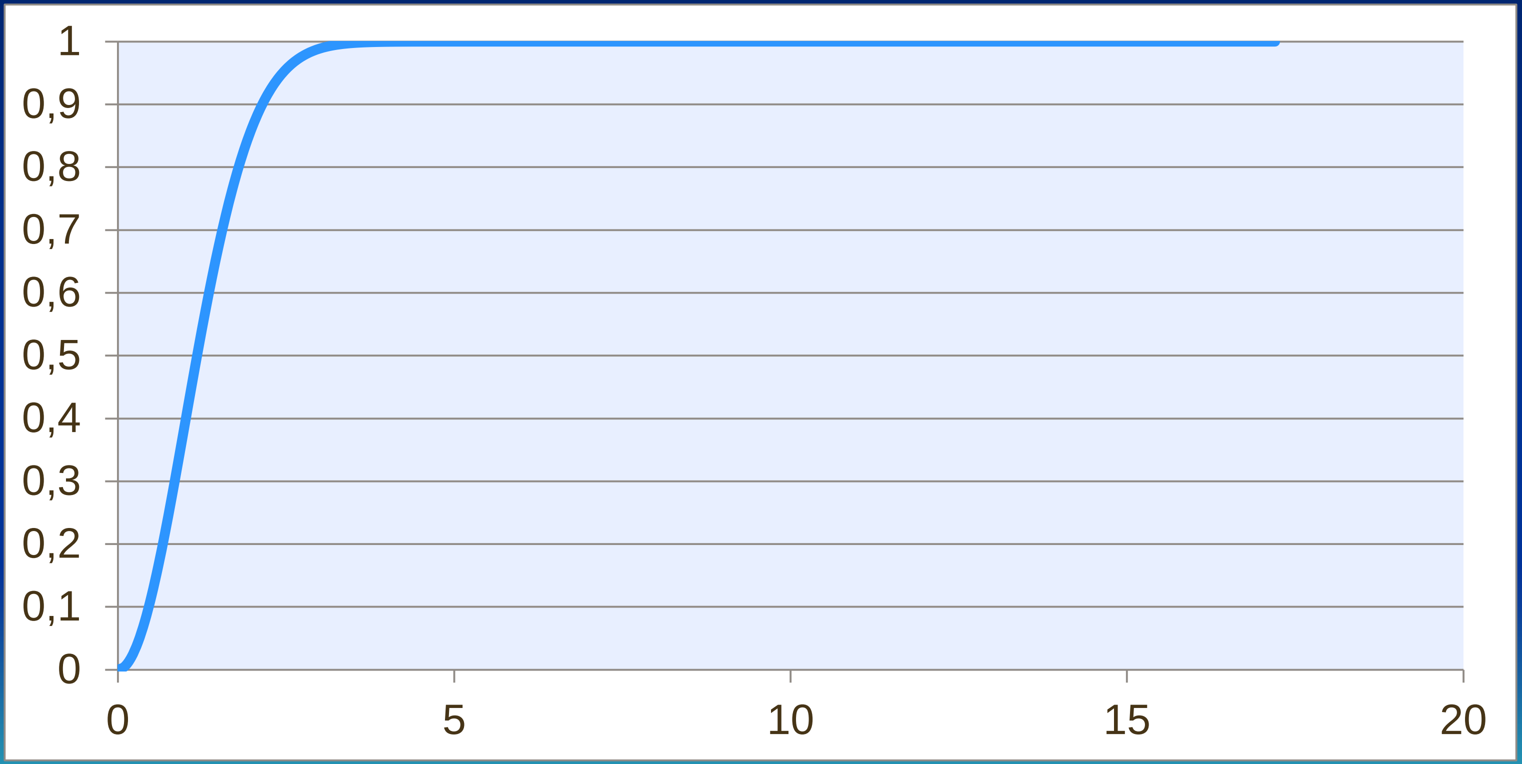
Calcolo della funzione di ripartizione

- F(t)=P(X<t)= $\int_0^t x e^{-\frac{x^2}{2}}\, dx$

$$\int_0^t x e^{-\frac{x^2}{2}}\, dx = 1 - e^{-\frac{t^2}{2}}$$

P(X>t) = 1- F(t)= exp(-t^2/2)

Rappresentazione grafica della funzione di ripartizione



Esercizio

Un'azienda produce occhiali utilizzando tre diversi macchinari. Il primo macchinario produce mediamente un paio di occhiali difettosi ogni 100, il secondo ogni 200, il terzo ogni 300. Gli occhiali vengono imballati in scatole identiche, contenenti 100 paia.

Ogni scatola contiene occhiali scelti a caso tra quelli prodotti da una sola delle tre macchine.

Si supponga che il primo macchinario abbia una produzione doppia rispetto agli altri due, cioè una scatola scelta a caso ha probabilità 1/2 di essere prodotta dal primo macchinario, 1/4 dal secondo e 1/4 dal terzo.

Quesiti

- Un ottico riceve una scatola con 100 paia di occhiali.

1. Qual è la probabilità che trovi almeno un paio di occhiali difettoso?
2. Se l'ottico trova esattamente due paia difettose, qual è la probabilità che gli occhiali siano stati prodotti dal primo macchinario?

Soluzione

- P(Di)=Prob di avere un paio di occhiali difettosi dalla macchina i   i=1,2, 3
- P(D1)=1/100   P(D2)=1/200 P(D3)=1/300
- Pi=Prob. che gli occhiali provengano dalla macchina i
- P1=1/2   P2=1/4   P3=1/4
- Richiesta: P(almeno un paio di occhiali difettoso nella scatola da 100 pezzi)?
- Oss. P(almeno un paio di occhiali difettoso)= 1-P(nessun paio di occhiali difettoso)=1-P(ND)

$P(D1)=1/100 \quad P(D2)=1/200 \quad P(D3)=1/300$

$P1=1/2 \qquad P2=1/4 \qquad P3=1/4$

Obiettivo calcolare P(ND) = 1-P(nessun paio di occhiali difettoso nella scatola da 100 pezzi)

$P(ND)=P(ND|1)\times P1+ P(ND|2)\times P2+ P(ND|3)\times P3$

- $P(ND|i)$ = prob nessuno paio difettoso tra i 100 se questi sono prodotti dalla macchina i
- $P(ND|1)= (1-1/100)^{100}$
- $P(ND|2)= (1-1/200)^{100}$
- $P(ND|3)= (1-1/300)^{100}$

Obiettivo calcolare P(ND)

- $P(ND|1)= (1-1/100)^{100}=0.366$
- $P(ND|2)= (1-1/200)^{100} = 0.6058$
- $P(ND|3)= (1-1/300)^{100}= 0.7161$

- P1=1/2 P2=1/4 P3=1/4

$P(ND)=P(ND|1)\times P1+ P(ND|2)\times P2+ P(ND|3)\times P3$

$= 0.366*1/2+0.6058*1/4 +0.7161*1/4=0.5135$

P(almeno un paio di occhiali difettoso)

- P(almeno un paio di occhiali difettoso nella scatola da 100 pezzi)=
  1-P(nessun paio di occhiali difettoso nella scatola da 100 pezzi)

  =1-P(ND)

  =1- 0.5135 = 0.4865

Se l'ottico trova esattamente due paia difettose, qual è la probabilità che gli occhiali siano stati prodotti dal primo macchinario?

- P(1|2D)= Prob che gli occhiali provengano dalla macchina 1 dato che nella scatola da 100 pezzi ce ne sono due difettosi

$$P(1|2D) = \frac{P(2D|1)\times P1}{P(2D|1)\times P1 + P(2D|2)\times P2 + P(2D|3)\times P3}$$

$$P(1|2D) = \frac{P(2D|1)\times P1}{P(2D|1)\times P1+P(2D|2)\times P2+P(2D|3)\times P3}$$

- P(D1)=1/100 P(D2)=1/200 P(D3)=1/300
- P1=1/2 P2=1/4 P3=1/4

Es. P(2D|1)P1=? $\binom{100}{2}\left(\frac{1}{100}\right)^2\left(1-\frac{1}{100}\right)^{98}\cdot\frac{1}{2}$

Es. P(2D|2)P2=? $\binom{100}{2}\left(\frac{1}{200}\right)^2\left(1-\frac{1}{200}\right)^{98}\cdot\frac{1}{4}$

$$= \frac{\binom{100}{2}\left(\frac{1}{100}\right)^2\left(1-\frac{1}{100}\right)^{98}\cdot\frac{1}{2}}{\binom{100}{2}\left(\left(\frac{1}{100}\right)^2\left(1-\frac{1}{100}\right)^{98}\cdot\frac{1}{2}+\left(\frac{1}{200}\right)^2\left(1-\frac{1}{200}\right)^{98}\cdot\frac{1}{4}+\left(\frac{1}{300}\right)^2\left(1-\frac{1}{300}\right)^{98}\cdot\frac{1}{4}\right)} =$$

$$= \frac{0.092432}{0.121275} = 0.762172$$

Esercizio

- Si lancia ripetutamente una coppia di dadi non truccati e si sommano i risultati.
- 1. Si calcoli la prob di ottenere un 7 come somma.
- 2. Si calcoli la prob che occorrano meno di 6 lanci per ottenere almeno un 7.
- 3. Si calcoli la prob che occorrano più di 6 lanci per ottenere almeno un 7

Prob. di ottenere un 7 come somma dei due dadi

Casi favorevoli :{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)}

- Prob. di ottenere un 7 =6/36=1/6

Si calcoli la prob che occorrano meno di 6 lanci per ottenere almeno un 7

- Prob di ottenere almeno un 7 in 5 lanci =
- 1-P(nessun sette in 5 lanci)
- $=1-(5/6)^5$

Si calcoli la prob che occorrano più di 6 lanci per ottenere almeno un 7

- Prob di non ottenere un 7 nei primi 6 lanci = P(nessun sette nei primi 6 lanci)
- $= (5/6)^6$

Esercizio

- Per 10 paesi dell'Unione Europea si è osservato il prezzo in euro di un litro di benzina (X) e il numero di veicoli pro capite circolanti (Y). Si conoscono i seguenti risultati relativi alle due variabili.

$$\sum_{i=1}^{10} x_i = 8.79 \qquad \sum_{i=1}^{10} y_i = 8.63$$

$$\sum_{i=1}^{10} x_i^2/10 = 0.77385 \qquad \sum_{i=1}^{10} y_i^2/10 = 0.7695$$

- Inoltre è noto che la devianza residua è pari a 0.01157.

Richieste

- Calcolare i parametri *a* e *b* della retta di regressione assumendo Y come variabile dipendente (in mancanza di altre informazioni fare opportune ipotesi, giustificandole, sul segno del coefficiente angolare).
- Commentare la bontà di adattamento del modello.
- Calcolare l'intervallo di confidenza di β al livello di confidenza del 95% e commentare i risultati ottenuti.

$$\sum_{i=1}^{10} x_i = 8.79 \qquad \sum_{i=1}^{10} y_i = 8.63$$
$$\sum_{i=1}^{10} x_i^2/10 = 0.77385 \qquad \sum_{i=1}^{10} y_i^2/10 = 0.7695$$

Calcolo bontà di adattamento

- DEV(X)=($[M_2(X)]^2$-$[M(X)]^2$)n

  =0.77385-$(8.79/10)^2$=0.01209
- DEV(E)= 0.01157
- DEV(Y)=($[M_2(Y)]^2$-$[M(Y)]^2$)n

  =[0.7695 -$(8.63/10)^2$]=0.024731

$$\delta = \frac{DEV(\hat{Y})}{DEV(Y)} = 1 - \frac{DEV(E)}{DEV(Y)}$$

δ=1-0.01157/0.024731=0.953

in mancanza di altre informazioni fare opportune ipotesi, giustificandole, sul segno del coefficiente angolare

$\delta=1-0.01157/0.24731=0.953$

Modello molto soddisfacente

$r_{xy}=- (0.953)^{0.5}$

Ip. Prendiamo il segno negativo poiché è lecito ipotizzare una relazione inversa tra costo della benzina e auto in circolazione

Calcolo del coefficiente angolare

$$b = \frac{COV(X,Y)}{VAR(X)} = r_{xy} \frac{\sigma_y}{\sigma_x}$$

$$a = \bar{y} - b\bar{x}$$

- b=-0.976 $(0.024731 / 0.001209)^{0.5}$ =-4.41
- a= 4.74

$$\sum_{i=1}^{10} x_i = 8.79 \qquad \sum_{i=1}^{10} y_i = 8.63$$

Calcolo dell'intervallo di confidenza al 95% di β, n=10

$$s = \sqrt{\frac{\sum e_i^2}{n-2}}$$

s = (0.01157/8)$^{0.5}$=0.038

$$s_{\hat{\beta}} = SE(\hat{\beta}) = \frac{s}{\sqrt{\sum (x_i - \bar{x})^2}}$$

= 0.038 / (0.01209)$^{0.5}$=0.3459

$$P\left\{\hat{\beta} - t(\alpha/2)SE(\hat{\beta}) \le \beta \le \hat{\beta} + t(\alpha/2)SE(\hat{\beta})\right\} = 1-\alpha$$

$$P\{-4.41 - 2{,}306 \times 0.3459 \le \beta \le -4.41 + 2.306 \times 0.3459\} = 0.95$$

$$P\{-5.2076 \le \beta \le -3.6124\} = 0.95$$