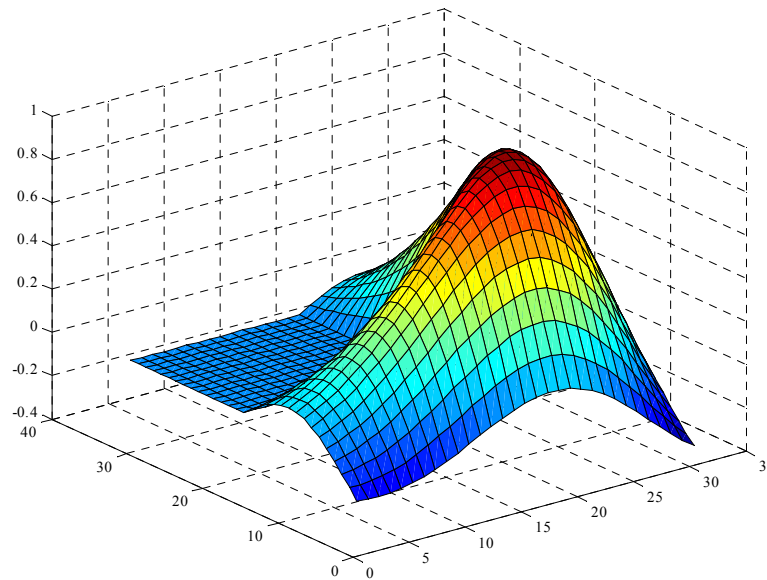


Dalmine, 28 aprile 2004

Esercitazioni di Statistica con Matlab

Dott. Orietta Nicolis

orietta.nicolis@unibg.it



V.c.d. di Poisson

$$P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda}, \quad x = 1, \dots, n$$

- Momenti

$$E(X) = \lambda$$

$$\text{Var}(X) = \lambda$$

Esempio

Un venditore di una grande concessionaria di automobili, riesce a vendere in media 5 automobili al giorno. Supponendo che il n° di auto vendute segua una distribuzione di Poisson, determinare:

- A) La distr. di prob. e di ripartizione del n. di auto vendute*
- B) La prob. che il n. delle auto vendute in un certo giorno sia almeno uguale a 3;*
- C) La prob. che il numero delle auto vendute sia compreso tra 4 e 6.*

Distribuzione di probabilità

$$P(x) = \frac{5^x}{x!} e^{-5}$$

`poisspdf(x, λ)`

- `x = 0:20;`
- `y = poisspdf(x, 5);`
- `ris = [x' y']`
- `bar(x, y), grid on`

Funzione di ripartizione

poisscdf(x, λ)

- `x = 0:20;`
- `y = poisscdf(x, 5);`
- `stairs(x,y)`
- `hold on`
- `plot(x,'o')`
- `hold off`

B) *La prob. che il n. auto sia $> 0 = 3$ è*

$$P(x \geq 3) = 1 - F(2) = 1 - \text{poisscdf}(2, 5)$$

C) *La prob. che il n. auto sia compreso tra 4 e 6 è*

$$\begin{aligned} P(4 \leq x \leq 6) &= F(6) - F(3) = \\ &= \text{poisscdf}(6, 5) - \text{poisscdf}(3, 5) \end{aligned}$$

Esempio

- *Supponendo di eseguire n prove indipendenti su un evento di probabilità p , determinare la probabilità di ottenere 3 volte l'evento se:*
 - A) $n=10$ e $p=0.3$;*
 - B) $n=100$ e $p=0.03$;*
 - C) $n=1000$ e $p=0.03$;*
 - D) $n = 10000$ e $p=0.003$*

Soluzione

Distribuzione binomiale $\rightarrow \text{Bin}(n,p)$

$$yA = \text{binopdf}(3, 10, 0.3)$$

$$yB = \text{binopdf}(3, 100, 0.03)$$

$$yC = \text{binopdf}(3, 1000, 0.003)$$

$$yD = \text{binopdf}(3, 10000, 0.0003)$$

Distribuzione di Poisson $\rightarrow \text{Po}(\lambda)$, con $\lambda = np = 3$

$$yD = \text{poisspdf}(3, 3)$$

V.c.d Uniforme

$$P(X = x) = \frac{1}{n}, \quad x = 1, \dots, n$$

Momenti:

$$E(X) = \frac{n+1}{2}$$

$$\text{Var}(X) = \frac{n^2 - 1}{12}$$

In Matlab:

`unidpdf(x,n)`, `unidcdf(x,n)` , `unidstat(n)`

Esempio

- *Descrivere (graficamente) con un'opportuna variabile casuale il punteggio ottenibile dal lancio di un dado.*

Soluzione

- *v.c. uniforme discreta con $n = 6$;*

```
n=6; x=1:n;
```

```
y =unidpdf(x, n)
```

```
subplot(2,1,1), bar(x,y)
```

```
yc=unidcdf(x, n)
```

```
subplot(2,1,2), stairs(x, yc)
```

```
hold on
```

```
plot(x, yc, 'o')
```

```
hold off
```

V.c.d. geometrica

$$P(X = x) = p \cdot (1 - p)^x, \quad x = 0, 1, \dots, \infty$$

Momenti:

$$E(X) = \frac{q}{p}$$

$$\text{Var}(X) = \frac{q}{p^2}$$

Esercizio

- *Si vuole testare un nuovo processo produttivo. Si sa che la probabilità di guasto è del 1%.*
 - a) *Descrivere la distribuzione di probabilità del tempo (in ore) finchè si verifica il guasto.*
 - b) *Qual è la probabilità che la macchina si fermi nella prima ora?*
 - c) *Quante ore occorre attendere affinché si abbia il guasto con probabilità del 99,9%.*

V.c.d. binomiale negativa

Bin⁻(r, p)

$$P(X = x) = \binom{r + x - 1}{x} p^r (1 - p)^x, \quad x = 0, 1, \dots$$

r = n° di successi ,

x = n° di insuccessi

In Matlab:

`nbinpdf(x, r, p)`

Esempio

- *Rappresentare graficamente le distribuzioni di probabilità e le funzioni di ripartizione con $p = 0.5$ e $r=2, 3, 4$.*
- *Confrontare la distribuzione di probabilità geometrica, di parametro $p = 0.2$, con quella binomiale negativa, di parametri $p = 0.2$ e $r = 1$.*

Distribuzioni di probabilità

- `x=0:10;`
- `y1=nbino.pdf(x, 2, 0.5);`
- `y2=nbino.pdf(x, 3, 0.5);`
- `y3=nbino.pdf(x, 4, 0.5);`
- `plot([y1' y2' y3'], '*')`
- `set(gca, 'XLim', [-0.5, 10.5])`

Funzione di ripartizione

- `x=0:10;`
- `y1=nbincdf(x, 2, 0.5);`
- `y2=nbincdf(x, 3, 0.5);`
- `y3=nbincdf(x, 4, 0.5);`
- `stairs(x, [y1' y2' y3'])`
- `set(gca, 'XLim', [-0.5, 10.5])`
- `hold on`
- `plot(x, [y1' y2' y3'], 'o'), hold off`

Confronto

- `x=0:10;`
- `subplot(2,1,1)`
- `y1=nbinpdf(x, 1, 0.2);`
- `plot(x, y1, '*')`
- `subplot(2,1,2)`
- `y2=geopdf(x, 0.2);`
- `plot(x, y2, '*')`

V.c.c. Normale

$N(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < +\infty$$

$z = \frac{x - \mu}{\sigma} \rightarrow$ La distribuzione Normale Standard $N(0,1)$

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, \quad -\infty < x < +\infty$$

... in Matlab

- Densità di probabilità:

`normpdf(x, mu, sigma);`

- Funzione di ripartizione

`normcdf(x, mu, sigma);`

- Momenti: `[m v]=normstat(mu, sigma);`

- F. inversa: `norminv(P, mu, sigma)`

- Generazione di numeri casuali:

`normrnd(mu, sigma, nr, nc) % da $N(\mu, \sigma)$`

`randn(nr, nc) % da $N(0,1)$`

Esempio 5

Confrontare le distribuzioni di probabilità e le funzioni di ripartizione delle seguenti variabili casuali continue Normali

$$X1 \sim N(0, 1),$$

$$X2 \sim N(4, 0.6),$$

$$X3 \sim N(2, 0.3),$$

$$X4 \sim N(-3, 2)$$

Distribuzioni di probabilità ...

```
x=-10:0.01:10;  
y1=normpdf(x, 0, 1);  
y2=normpdf(x, 4, 0.6);  
y3=normpdf(x, 2, 0.3);  
y4=normpdf(x, -3, 2);  
figure(1)  
plot(x,[y1' y2' y3' y4'])  
title('Distribuzioni Normali')  
legend('E(X)=0, V(X)=1', 'E(X)=4, V(X)=0.6',  
      'E(X)=2, V(X)=0.3', 'E(X)=-3, V(X)=2')
```

... funzioni di ripartizione

Figure(2)

```
yc1=normcdf(x, 0, 1);
```

```
yc2=normcdf(x, 4, 0.6);
```

```
yc3=normcdf(x, 2, 0.3);
```

```
yc4=normcdf(x, -3, 2);
```

```
plot(x,[yc1' yc2' yc3' yc4'])
```

```
title('Funzioni di ripartizione')
```

Esempio (P.I. 15/04/03)

La lunghezza di certe barrette d'acciaio prodotte in serie da una macchina è una v.c. normale di media 4 cm e varianza 0.04 cm^2 .

- a) Determinare la probabilità che lo scarto tra la lunghezza osservata ed il valor medio sia minore di 0.1 cm. e disegnare il grafico corrispondente.
- b) Se decidiamo di scartare il 5% delle barrette che hanno lunghezza minore e il 5% delle barrette che hanno lunghezza maggiore, determinare la lunghezza minima e massima dei pezzi accettati.
- c) Qual'è la probabilità che su 50 pezzi prodotti almeno 3 abbiano una lunghezza maggiore di 4.3.

....Soluzione

```
mu=4;
sigma2=0.04;
sigma=sigma2^0.5;
prob_punto_a=normspec([-0.1/sigma 0.1/sigma]);
xmin=norminv(0.05, mu, sigma);
normspec([-inf xmin], mu, sigma)
xmax=norminv(0.95, mu, sigma);
normspec([xmax inf], mu, sigma)
prob=1-normcdf(4.3, mu, sigma);
punto_c=1-binocdf(2, 50, prob);
sol=[prob_punto_a; xmin_punto_b1;
      xmax_punto_b2; prob; punto_c]
```

Esercizio

Sia X una v.c. con media $\mu=3$ e s.q.m $\sigma=4$.

- a) Si trovino $P(X>11)$.
- b) $P(2<X<7)$ e disegnare il grafico.
- c) Supponendo che la variabile X esprima il diametro di certi dischi metallici, determinare a quale diametro corrisponde il 2% dei dischi più grandi.

[tempo 5 min.]

Distribuzioni che derivano da quella normale

- La distribuzione chi-quadro
- La distribuzione t-Student
- La distribuzione F di Fisher

Distribuzione *chi-quadrato*

Se Z_1, Z_2, \dots, Z_k sono v.c. normali standard e indipendenti, allora la somma dei loro quadrati

$$X = Z_1^2 + Z_2^2 + \dots + Z_k^2$$

ha una distribuzione χ^2 con k g.d.l.

$$X \longrightarrow \chi_{(k)}^2$$

- È un caso particolare della distribuzione GAMMA che si ottiene se si pone

$$\lambda = \frac{1}{2} \quad \text{e} \quad r = \frac{k}{2},$$

Momenti:

$$E(X) = k, \quad \text{Var}(X) = 2 \cdot k$$

N.B. Per $k > 100$ si può utilizzare l'approssimazione

$$\sqrt{2\chi_{(k)}^2} - \sqrt{2k - 1} \xrightarrow[k \rightarrow \infty]{d} N(0,1)$$

In Matlab:

`chi2pdf(x,k),` `chi2pdf(x,k),`
`chi2stat(k),`

Esempio

Confrontare le funzioni di densità e di ripartizione con $k = 2, 4, 6, 8$.

```
x=0:.01:20;  
y1=chi2pdf(x, 2);  
y2=chi2pdf(x, 4);  
y3=chi2pdf(x, 6);  
y4=chi2pdf(x, 8);  
plot(x,y1, x, y2, x, y3, x, y4)  
title('Funzione di densità')  
legend('k=2', 'k=4', 'k=6', 'k=8')
```

(idem per la funzione di ripartizione!)

Esempio

Si determini $P(X \leq 30)$ quando X è una v.c. chi-quadro con 26 g.d.l.

...in matlab

```
Y=chi2cdf(30, 26)
```


Esempio

Si trovi quanto vale

$$P(\chi^2_{0.05,15})$$

...in matlab

`chi2inv(0.05, 15)`

Esercizio

Sia X una v.c. chi-quadro con 10 g.d.l.

- a) Disegnare la distribuzione di probabilità
- b) Si determini $P(X < 5)$;
- c) Si determini a quale numero corrisponde il 5% dei valori più bassi

[tempo 5 min.]

Distribuzione *t-Student*

Siano Z e C due v.c. indipendenti tali che

$$Z \sim N(0,1) \quad \text{e} \quad C \sim \chi^2_{(k)}$$

allora la v.c.

$$T = \frac{Z}{\sqrt{\frac{C}{N}}}$$

ha distribuzione t con k g.d.l.

$$T_k \sim t_{(k)}$$

Momenti:

$$E(X) = 0, \quad \text{Var}(X) = \frac{k}{k-2}$$

In Matlab:

tpdf(x,k), tcdf(x,k), tstat(r,λ),

Esempio

Confrontare le funzioni di densità e di ripartizione con $k = 1, 5, 10, 50$.

```
x=-10:.01:10;  
y1=tpdf(x, 1);  
y2=tpdf(x, 5);  
y3=tpdf(x, 10);  
y4=tpdf(x, 50);  
plot(x,y1, x, y2, x, y3, x, y4)  
title('Funzione di densità')  
legend('k=1', 'k=5', 'k=10', 'k=50')
```

(idem per la funzione di ripartizione!)

Esempio

Sia T una v.c. t-Student con 12 g.d.l.; si trovino:

a) $P(T > 1.4)$

b) $t_{0.025, 9}$

...in matlab

`1-tcdf(1.4, 12)`

`tinvs(0.025, 9)`

Esercizio

Sia T una v.c. t-Student con 8 g.d.l..

a) Disegnare la funzione cumulata di probabilità

b) Si trovi $P(T > 2)$

b) Si determini a quale numero corrisponde il 5% dei valori più bassi

[tempo 5 min.]

Distribuzione F

Siano U e V due v.c. indipendenti tali che

$$U \sim \chi^2_{(k_1)} \quad \text{e} \quad V \sim \chi^2_{(k_2)}$$

Allora la v.c.

$$F = \frac{\frac{U}{k_1}}{\frac{V}{k_2}}$$

Ha distribuzione F con k_1 e k_2 g.d.l

$$F \sim F(k_1, k_2)$$

Esempio

Sia X una v.c. F con 6 e 14 g.d.l.; si trovino:

a) $P(F \leq 1.5)$

b) $F_{5,12}$

...in matlab

`fcdf(1.5, 6, 14)`

`finv(0.025, 5, 12)`

Esempi Riepilogativi

Esempio 1

Una macchina è calibrata per produrre dischi metallici di raggio pari a 3 cm. Il raggio effettivo di un disco segue la distribuzione normale con media pari a 3 e varianza pari a 0.04.

- a) Rappresentare graficamente la distribuzione di probabilità e la funzione di ripartizione del raggio dei dischi.*
- b) Calcolare la % di dischi con raggio inferiore a 2.8 cm.*
- c) Determinare la % dei dischi con raggio maggiore di 3.25 cm.*
- d) Calcolare la % dei dischi con raggio compreso tra 2.6 e 3.4 cm*

a) Rappresentazione grafica

```
x=2:.01:4;  
mu=3;  
sigma=sqrt(0.04);  
y=normpdf(x, mu, sigma);  
yc=normcdf(x, mu, sigma);  
subplot(1,2,1)  
plot(x,y), title('Distribuzione N(3,0.04)')  
subplot(1,2,2)  
plot(x,yc), title('Funzione di ripartizione')
```

b) $P(x \leq 2.8) = F(2.8)$

```
yb=normcdf(2.8, mu, sigma);  
normspec([-Inf 2.8], mu, sigma);
```

c) $P(x > 3.25) = 1 - F(3.25)$

```
yc=1-normcdf(3.25, mu, sigma);  
normspec([3.25 +Inf], mu, sigma);
```

b) $P(2.6 < x \leq 3.4) = F(2.8)$

```
yc=normcdf(3.4, mu, sigma)- normcdf(2.6, mu,  
    sigma);  
normspec([2.6 3.4], mu, sigma);
```

Esempio 2

Una ditta ha sottoposto ad un test di durata 5000 batterie da essa costruite. La durata media è risultata pari a 100 ore e la deviazione standard pari a 6 ore. Supponendo che la distribuzione delle batterie secondo la durata sia normale, si determini:

- a) La distribuzione di probabilità della durata delle batterie.*
- b) Il numero di batterie con durata minore di 80 ore.*
- c) La durata minima del 10% delle batterie che durano di più.*

a) Distribuzione di probabilità

```
x=80:.01:120;  
mu = 100; sigma = 6;  
y=normpdf(x, mu, sigma);  
yc=normcdf(x, mu, sigma);  
subplot(1,2,1)  
plot(x,y), title('Distribuzione N(100,6)')  
subplot(1,2,2)  
plot(x,yc), title('Funzione di ripartizione')
```

*b) n. di batterie con durata minore di 90 ore.
→ $N.batterie * F(90)$*

```
yb=normcdf(90, mu, sigma);  
normspec([-Inf 90], mu, sigma);  
n_batt=yb*5000
```

c) Durata minima del 10% che durano di più.

```
yc=norminv(0.90, mu, sigma)  
normspec([yc +Inf], mu, sigma);
```

Esempio 3

Un prodotto si ottiene dall'assemblaggio di 3 componenti. La lunghezza complessiva del prodotto Y , è uguale alla somma delle lunghezze X_1 , X_2 , X_3 delle sue componenti. Data la variabilità del processo si può assumere che

$$X_1 \sim N(2, 0.01), \quad X_2 \sim N(4, 0.02),$$

$$X_3 \sim N(3, 0.02),$$

- a) Rappresentare le distribuzioni di probabilità delle tre componenti e di $Y=X_1+X_2+X_3$*
- b) Determinare la probabilità che la lunghezza del singolo pezzo prodotto soddisfi lo standard qualitativo prefissato, 9 ± 0.25*


```

a) x=0:.01:11;
mu1= 2; mu2= 4; mu3= 3; mut=mu1+mu2+mu3;
sigma1=sqrt(0.01); sigma2=sqrt(0.02);
    sigma3=sqrt(0.02); sigmat=sqrt(0.01+0.02+0.02);
y1=normpdf(x, mu1, sigma1);
y2=normpdf(x, mu2, sigma2);
y3=normpdf(x, mu3, sigma3);
yt=normpdf(x, mut, sigmat);
plot(x,y1, x, y2, x,y3, x, yt)
legend('N(2,0.01)', 'N(4,0.02)', 'N(3,0.02)', 'yt')
title('Distribuzioni di probabilità')
b) p=normcdf(9+0.25, mut, sigmat)-normcdf(9-0.25,
    mut, sigmat),
normspec([9.25 8.75], mut, sigmat);

```

Esercizi riepilogativi

Esercizio 1

Le sfere d'acciaio prodotte da una fabbrica hanno un diametro che segue la distribuzione normale. Si supponga che il 20% delle sfere abbia un diametro inferiore a 1.2 cm e che il 10% abbia un diametro maggiore di 1.4 cm.

- A) Determinare il valore atteso del diametro di una sfera.*
- B) Dare una misura dell'inaccuratezza del processo (varianza).*
- C) Sulla base dei risultati precedenti calcolare la probabilità che il diametro di una sfera sia maggiore di 1.5.*

Esercizio 2

In una certa fabbrica il responsabile del controllo di qualità sa, per esperienza acquisita, che il 4.5% della produzione presenta difetti. Qual è la probabilità che in un nuovo lotto vi siano non più di 24 pezzi difettosi? (Usare l'approssimazione normale).

Esercizio 3

Una ditta produttrice di televisori, sostiene, in base alla passata esperienza, che i suoi prodotti hanno una durata media pari a 10 anni e una deviazione standard pari a 2. La ditta ripara gratuitamente le lavatrici componenti che si rompono durante la garanzia.. Assumendo che la distribuzione secondo la durata segua una distribuzione normale, quale deve essere la durata della garanzia, affinché l'azienda sia chiamata a riparare solo il 3% delle lavatrici vendute?