Laboratorio di ST1
Lezione 3

Claudia Abundo

Dipartimento di Matematica
Università degli Studi Roma Tre

19 Marzo 2010

Distribuzioni Statistiche

- Discrete
- Continue

Comandi principali

dNOME() Funzione di densità o distribuzione di probabilità per variabili discrete
pNOME() Funzione di ripartizione
qNOME() Quantili
rNOME() Generatore di numeri pseudo-casuali
Ogni distribuzione ha dei suoi parametri: richiamare l'help per una corretta definizione di questi

ESEMPIO: distribuzione Binomiale

```
dbinom(size, prob)
pbinom(size, prob)
qbinom(size, prob)
rbinom(size, prob)
```

Densità, funzione di distribuzione, quantili e funzione generatrice di numeri casuali per una funzione di distribuzione binomiale di parametri numerosità (size) e probabilità (prob)

Distribuzioni Discrete

- Binomiale
- Poisson

Distribuzioni Continue

- Uniforme
- Normale

Discrete 1: Binomiale Bin(p, n)

$$\binom{n}{x} p^x (1-p)^{n-x}$$

Valore Atteso: np Varianza: $np(1-p)$

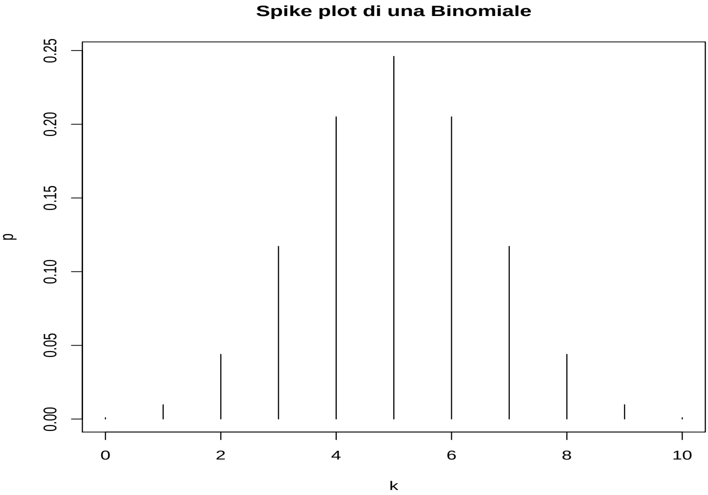
Ipotizziamo di avere una popolazione che si distribuisce come una variabile aleatoria $X \sim Bin(n, p)$ ossia che le n variabili aleatorie $X_1, X_2, \dots, X_n$ hanno distribuzione di probabilità discreta

$$f_{X_i}(x_i|n, p) = \binom{n}{x_i} p^{x_i} (1-p)^{n-x_i}$$

e tale che $E(X_i) = np$ e $var(X_i) = np(1-p)$

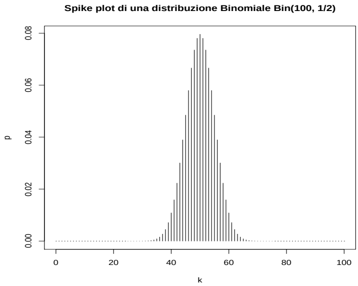
Consideriamo la probabilità associata alla sequenza da 0 a 10 da una Bin(10, 1/2) e disegnamo lo Spike plot

```
BINO = dbinom(0:10, size = 10, prob = 1/2)
plot(0:10, BINO, type = "h",
main= "Spike plot di una distribuzione Binomiale Bin(10, 1/2)",
xlab= "k", ylab = "p")
```


Spike plot di una Binomiale
p
0.00
0.05
0.10
0.15
0.20
0.25
0
2
4
6
8
10
k

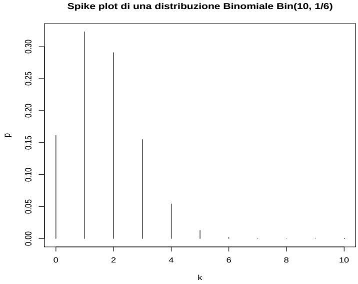
Cambiamo i parametri n e p _1

```
BINO = dbinom(0:100, size = 100, prob = 1/2)

plot(0:100, BINO, type = "h", main= "Spike plot di una distribuzione
Binomiale Bin(100, 1/2)", xlab= "k", ylab = "p")
```



Cambiamo i parametri n e p _2

```
BINO = dbinom(0:10, size = 10, prob = 1/6)

plot(0:10, BINO, type = "h", main= "Spike plot di una distribuzione
Binomiale Bin(10, 1/6)", xlab= "k", ylab = "p")
```



Discrete 2: Poisson

o anche “legge degli eventi rari”

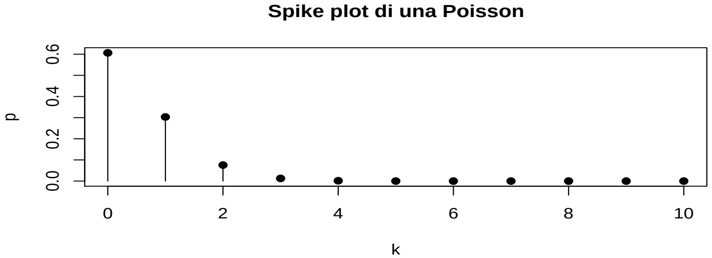
$$\lambda^x \frac{exp(-\lambda)}{x!}$$

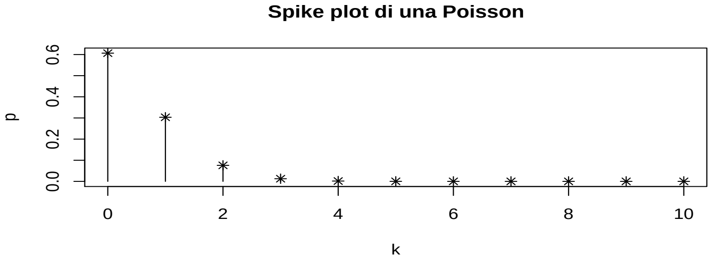
Valore Atteso: λ = Varianza: λ
Ora si usa dpois(x, lambda):

```
par(mfrow = c(2, 1))

POI = dpois(0:10, lambda=1/2)
plot(0:10, POI, type = "h", main=
"Spike plot di una Poisson", xlab="k", ylab = "p")

points(0:10,POI,pch = 16, cex = 1)
```

oppure possiamo cambiare i pallini in “stars”:

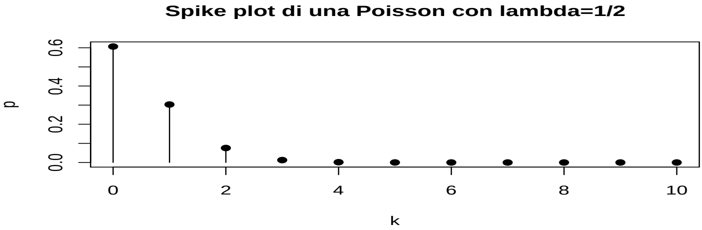
```
POI2 = dpois(0:10, lambda=1/2)
plot(0:10, POI2, type = "h", main=
"Spike plot di una Poisson", xlab="k", ylab = "p")

points(0:10,POI2,pch = 8,cex = 1)
```

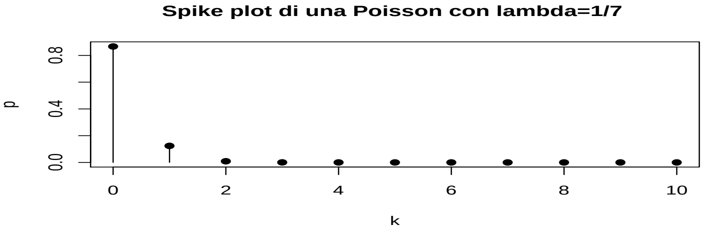

Spike plot di una Poisson
p
0.0
0.2
0.4
0.6
0
2
4
6
8
10
k


Spike plot di una Poisson
p
0.0
0.2
0.4
0.6
0
2
4
6
8
10
k

Cambiamo il parametro λ

```
par(mfrow = c(2, 1))

POI = dpois(0:10, lambda=1/2)
plot(0:10, POI, type = "h", main=
"Spike plot di una Poisson con lambda=1/2", xlab="k", ylab = "p")
points(0:10,POI,pch = 16, cex = 1)

POI2 = dpois(0:10, lambda=1/7)
plot(0:10, POI2, type = "h", main=
"Spike plot di una Poisson con lambda=1/7 ", xlab="k", ylab = "p")
points(0:10,POI2,pch = 16, cex = 1)
```

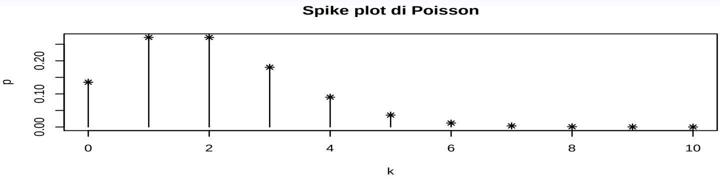

Spike plot di una Poisson con lambda=1/2
p
0.0
0.2
0.4
0.6
0
2
4
6
8
10
k

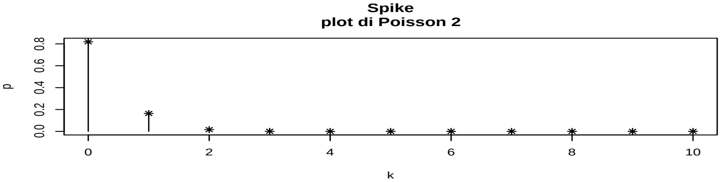

Spike plot di una Poisson con lambda=1/7
p
0.0
0.4
0.8
0
2
4
6
8
10
k

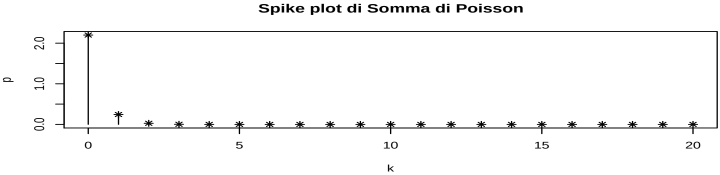
Somma di Poisson

La somma di variabili aleatorie *indipendenti* con distribuzioni di Poisson di parametri λ_1 e λ_2 è una variabile aleatoria con distribuzione di Poisson di parametro $\lambda = \lambda_1 + \lambda_2$

```
par(mfrow = c(3, 1))

POI = dpois(0:10, lambda=2)
plot(0:10,POI,type = "h", main= "Spike
plot di Poisson", xlab="k", ylab ="p")

points(0:10,POI,pch = 8,cex =1)

POI2 = dpois(0:10, lambda=1/5)
plot(0:10, POI2, type="h",main=
"Spike plot di Poisson 2", xlab="k", ylab = "p")

points(0:10,POI2,pch=8,cex= 1)

POI3=dexp(0:20, rate = 2+1/5)
plot(0:20,poi3,type = "h", main=
"Spike plot di Somma di Poisson", xlab="k", ylab ="p")

points(0:20,poi3,pch = 8,cex = 1)
```


Spike plot di Poisson
p
0.00
0.10
0.20
0
2
4
6
8
10
k


Spike
plot di Poisson 2
p
0.0
0.2
0.4
0.6
0.8
0
2
4
6
8
10
k

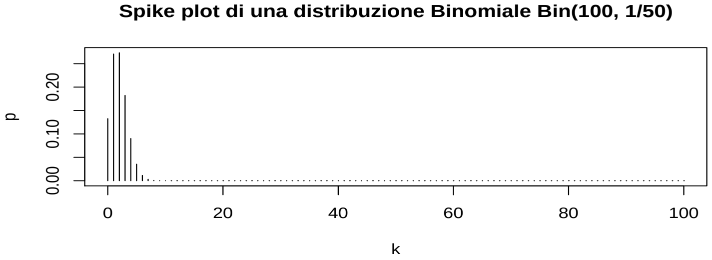

Spike plot di Somma di Poisson
p
0.0
1.0
2.0
0
5
10
15
20
k

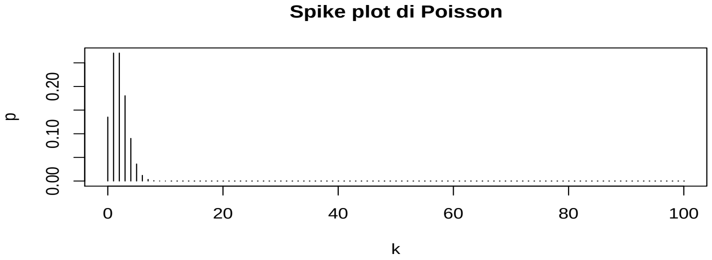
Binomiale $\longrightarrow$ Poisson

Se n èmolto grande (~ 50) e p molto piccolo, tale che $np < 10$ e $p(1-p) \sim p$, allora la binomiale può essere approssimata con una distribuzione di Poisson con parametro $\lambda = np$

ESEMPIO: $n = 100 \;\; p = 1/50 \;\; np = 2$

```
par(mfrow = c(2, 1))

BINO = dbinom(0:100, size = 100, prob = 1/50)
plot(0:100, BINO, type
= "h", main= "Spike plot di una distribuzione Binomiale Bin(100,
1/50)", xlab= "k", ylab = "p")

POI = dpois(0:100, lambda=2)
plot(0:100,POI,type = "h", main= "Spike
plot di Poisson", xlab="k", ylab ="p")
```


Spike plot di una distribuzione Binomiale Bin(100, 1/50)
p
0.00
0.10
0.20
0
20
40
60
80
100
k


Spike plot di Poisson
p
0.00
0.10
0.20
0
20
40
60
80
100
k

Continue 1: Uniforme U[a,b]

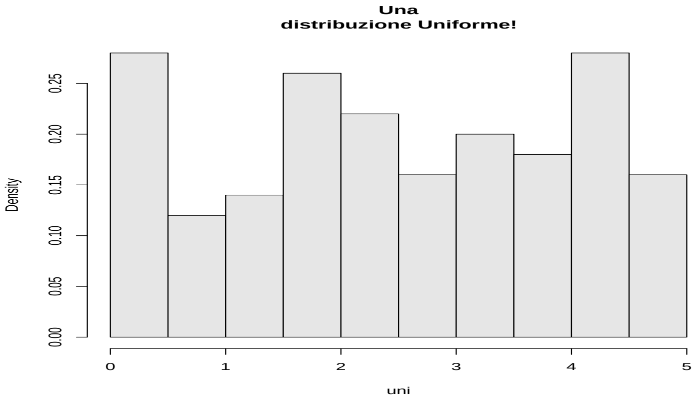
$$\frac{1}{b-a}$$

$a \leq x \leq b$, è nulla per $x < a$ e $x > b$
Ha due parametri denominati “a” e “b”, che costituiscono rispettivamente il minimo ed il massimo.
Valore atteso: $\frac{(a+b)}{2}$ Varianza: $\frac{(b-a)^2}{12}$

Simuliamo 100 valori da una Uniforme in (0, 5)

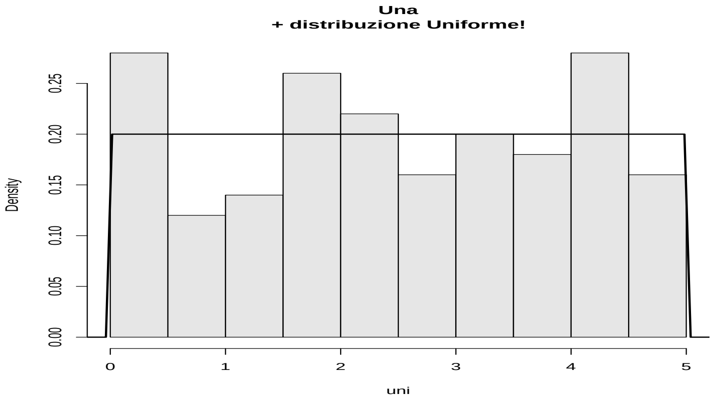
Costruiamo l’istogramma corrispondente

```
uni = runif(100,min = 0,max = 5)
uni
hist(uni, prob=TRUE, main= "Una
distribuzione Uniforme!", col=gray(0.9))
```


Una
distribuzione Uniforme!
Density
0.00
0.05
0.10
0.15
0.20
0.25
0
1
2
3
4
5
uni

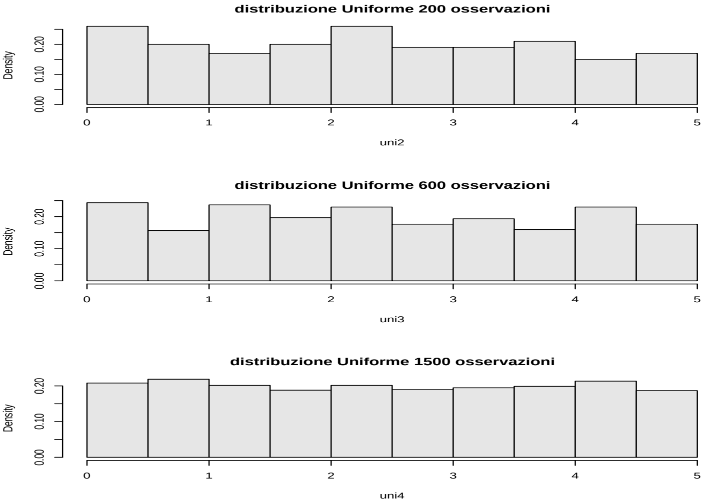
Con il comando “curve” disegnamo la curva Inserendo “add=TRUE” sovrapponiamo la curva all’istogramma

```
plot.new()
box()
hist(uni, prob=TRUE, main="Una distribuzione Uniforme!",
col=gray(0.9))

curve(dunif(x, min=0, max=5), lwd=2, add=TRUE)
```


Una
+ distribuzione Uniforme!
Density
0.00
0.05
0.10
0.15
0.20
0.25
0
1
2
3
4
5
uni

Cambiamo il numero di osservazioni

```
par(mfrow = c(3, 1))

uni2 = runif(200,min = 0,max = 5)
uni2
hist(uni2, prob=TRUE, main=
"distribuzione Uniforme 200 osservazioni", col=gray(0.9))

uni3 = runif(600,min = 0,max = 5)
uni3
hist(uni3, prob=TRUE, main=
"distribuzione Uniforme 600 osservazioni", col=gray(0.9))

uni4 = runif(1500,min = 0,max = 5)
uni4
hist(uni4, prob=TRUE, main=
"distribuzione Uniforme 1500 osservazioni", col=gray(0.9))
```

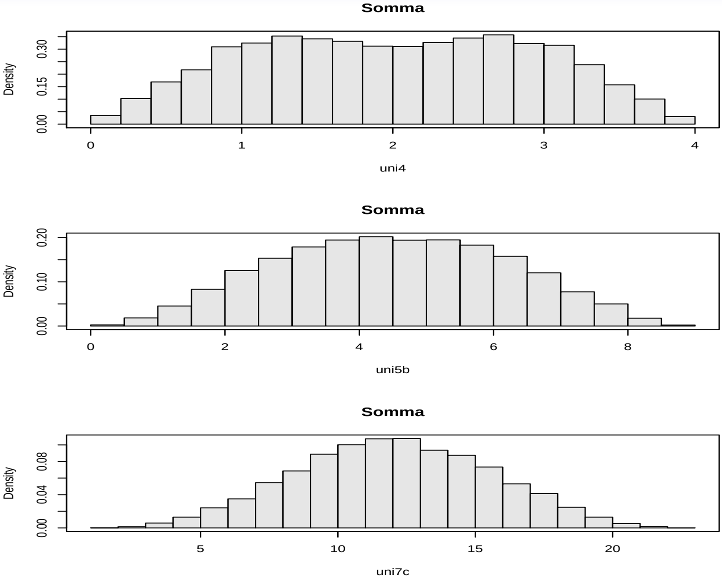

distribuzione Uniforme 200 osservazioni
Density
0.00
0.10
0.20
0
1
2
3
4
5
uni2
distribuzione Uniforme 600 osservazioni
Density
0.00
0.10
0.20
0
1
2
3
4
5
uni3
distribuzione Uniforme 1500 osservazioni
Density
0.00
0.10
0.20
0
1
2
3
4
5
uni4

Somma di Uniformi_2: Teorema del limite centrale

Attenzione: la somma di Uniformi non è un'Uniforme!

```
par(mfrow = c(3, 1))
uni2 = runif(10000,min = 0,max = 1)
uni3 =
runif(10000,min = 0,max = 3)
uni4=uni2+ uni3 hist(uni4, prob=TRUE,main="Somma", col=gray(0.9))
box()

uni2b = runif(10000,min = 0,max = 1)
uni3b = runif(10000,min = 0,max = 3)
uni4b = runif(10000,min = 0,max = 5)
uni5b=uni2b+ uni3b + uni4b
box()
hist(uni5b, prob=TRUE, main="Somma", col=gray(0.9))
```

```
uni2c = runif(10000,min = 0,max = 1)
uni3c = runif(10000,min = 0,max = 3)
uni4c = runif(10000,min = 0,max = 5)
uni5c = runif(10000,min =0,max = 7)
uni6c = runif(10000,min = 0,max = 8)

uni7c=uni2c+ uni3c + uni4c + uni5c + uni6c

box()
hist(uni7c, prob=TRUE, main="Somma",col=gray(0.9))

box()
```



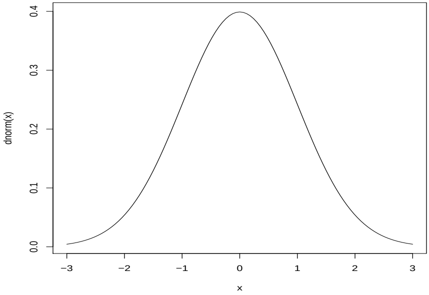
Somma
Density
0.00
0.15
0.30
0
1
2
3
4
uni4
Somma
Density
0.00
0.10
0.20
0
2
4
6
8
uni5b
Somma
Density
0.00
0.04
0.08
5
10
15
20
uni7c

Continue 2: Normale $N(\mu, \sigma^2)$

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Disegnamo i valori di una curva Normale tra -3 e 3
(Una Normale varia da $-\infty$ a ∞)
con il comando “curve”

```
curve(dnorm(x),-3,3, axes = TRUE, ylab = "", xlab = "", ylim =
c(0,.5))
```



Evidenziare le code

Evidenziamo la coda di SINISTRA

```
vals = seq(-3,-1, length = 100)
x =c(-3, vals, -1, -3)
y = c(0, dnorm(vals),0, 0)
polygon(x,y, density= 20, angle = 45)
```

Evidenziamo la coda di DESTRA

```
vals = seq(1,3,length = 100)
 x = c(1, vals, 3, 1)
 y = c(min(y), dnorm(vals),min(y), min(y))
 polygon(x,y,density = 20,angle = 45)
```