# R - Esercitazione 2

Lorenzo Di Biagio
dibiagio@mat.uniroma3.it

Università Roma Tre

14 Ottobre 2013

# Quantili (1)

Sia $0 \leq p \leq 1$.

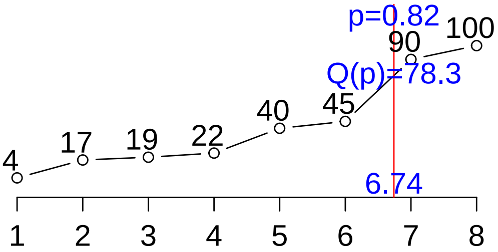
Il quantile $p$-esimo di una variabile causale $X$ è il più piccolo numero $\xi$ per cui $F_X(\xi) \geq p$.

Dato un insieme finito di valori, ordinati in modo non decrescente, il quantile $p$-esimo è un valore $\xi$ tale che la frazione di osservazioni inferiori o uguali a $\xi$ sia almeno $p$, mentre la frazione di osservazioni maggiori o uguali a $\xi$ non sia inferiore a $1 - p$.

$\xi$ in genere non è definito univocamente. Noi adotteremo la convenzione di R, i.e., il valore della funzione `quantile(x,p)` con parametro `type` implicitamente preimpostato a 7.
Mediana: quantile con p=0.5
Quartili: quantili con p=0.25,0.5,0.75
Percentili: quantili con p=0.01,0.02,. . . , 0.99

# Quantili (2)

Siano $x_1, x_2, \ldots, x_N$, $N$ osservazioni ordinate in modo non decrescente.
Allora $Q(p) = (1-\gamma)x_j + \gamma x_{j+1}$ dove:
$j = [1 + p(N-1)]$ ($[\cdot]$: parte intera inferiore)
$\gamma = \{1 + p(N-1)\}$ ($\{\cdot\}$: parte frazionaria)
E.g.:



# Matrici (1)

Le matrici in R si generano con `matrix` applicato ad un vettore e specificando il numero di righe (`nrow`) e il numero di colonne (`ncol`) oppure fissando le dimensioni di un oggetto con `dim`.

```
> matrix(1:10,2,5)
```

crea la matrice $\begin{pmatrix} 1 & 3 & 5 & 7 & 9 \\ 2 & 4 & 6 & 8 & 10 \end{pmatrix}$

Se si vuole che gli elementi del vettore siano disposti per riga, utilizzare l'opzione `byrow=T`.

`rbind()`, `cbind()`: fondono matrici per righe o per colonne.
`det()`, `solve()` : calcolano determinante e inverso di una matrice quadrata.
`t()` genera la matrice trasposta.
`diag()` genera matrici diagonali o estrae il vettore della diagonale principale.
`%*%` prodotto righe per colonne

# Matrici (2)

## Esercizio 1

1. Si definisca `A <- matrix(1:5,nrow=3,ncol=3)`. Cosa si ottiene con l'estrazione `A[2,3]` ? Ipotizzarlo e poi verificarlo con R.
2. Verificare che il seguente sistema lineare ha una e una sola soluzione e calcolarla: $\begin{cases} 3X+2Y+Z & = & 8 \\ X+7Y & = & 12 \\ 5Y-3Z & = & 3 \end{cases}$
3. Verificare che la matrice $\begin{pmatrix} \cos(3\pi/5) & -\sin(3\pi/5) \\ \sin(3\pi/5) & \cos(3\pi/5) \end{pmatrix}$ è ortogonale.

# Distribuzioni di probabilità (1)

In R sono già implementate le principali distribuzioni di probabilità.
Ad esempio:

- binomiale: `binom`
- poisson: `pois`
- normale: `norm`
- uniforme: `unif`
- chi quadrato: `chisq`
- t di Student: `t`

Per ognuna di esse si può ottenere la densità di probabilità (`d`), la funzione di ripartizione (`p`), i quantili (`q`), generazione di numeri casuali (`r`).

# Distribuzioni di probabilità (2)

Una variabile aleatoria $X$ si dice bernoulliana di parametro $p$ se assume il valore 1 con probabilità $p$ e il valore 0 con probabilità $1 - p$.

Una variabile aleatoria $X$ si dice binomiale di parametri $p$ e $n$ se è la somma di $n$ variabili aleatorie i.i.d bernoulliane di parametro $p$. $P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$.

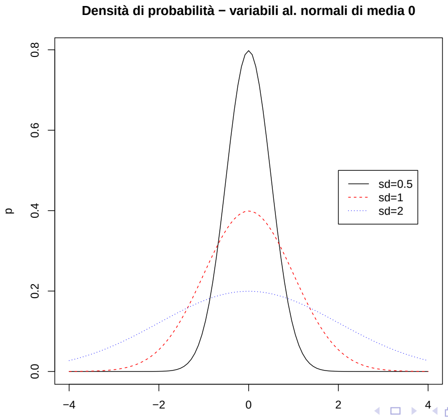
## Esercizio 2

1. Creare in un'unica pagina quattro grafici con le densità di probabilità di quattro binomiali di parametri $n = 10$ e $p = 1/100, 1/10, 1/2, 99/100$.
2. Inserire i quattro grafici in un unico grafico: le linee della densità saranno distinguibili dal colore e dalla legenda.

# Distribuzioni di probabilità (3)

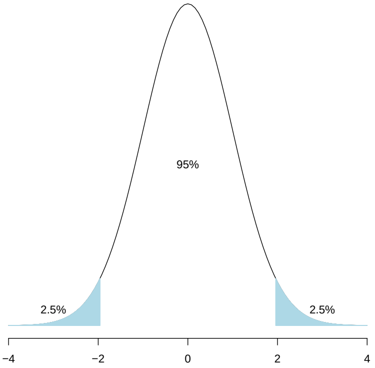
Una variabile aleatoria $X$ si dice normale di media $\mu$ e varianza $\sigma^2$ se la sua densità è data da

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



# Distribuzioni di probabilità (4)

Data la densità di probabilità di una normale di media 0 e varianza 1, evidenziamo la coda sinistra e la coda destra in modo che l'area centrale sia 0.95.



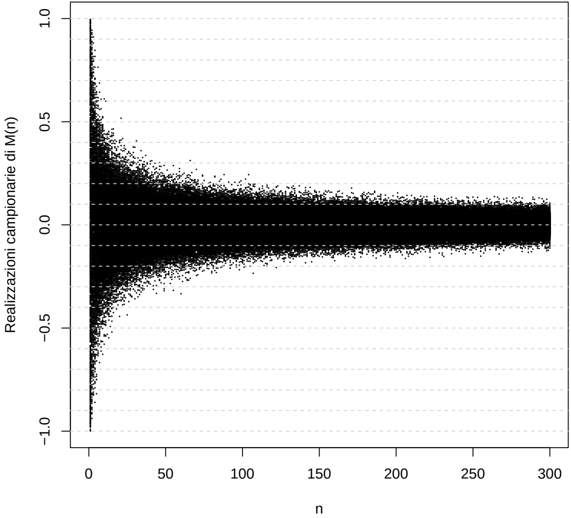
# La legge debole dei grandi numeri (1)

Sia $f(\cdot)$ una densità con media $\mu$ e varianza finita $\sigma^2$. Sia $M(n) = \overline{X}_n$ la media campionaria di un campione casuale di ampiezza $n$ da $f(\cdot)$. Sia $\epsilon > 0$. Allora

$$P(|\overline{X}_n - \mu| < \epsilon) \overset{n \to +\infty}{\longrightarrow} 1.$$

Si vuole simulare la legge dei grandi numeri considerando la densità uniforme tra -1 e 1: $f(x) = \frac{1}{2} I_{[-1,1]}(x)$. Per ogni $n = 1, \ldots, 300$ consideriamo 1000 realizzazioni campionarie di $\overline{X}_n$, indicando i punti su un grafico. Ci aspettiamo che più è alto $n$ più i punti si addensino attorno a $\mu = 0$.

# La legge debole dei grandi numeri (2)



# Il teorema limite centrale (1)

Sia $f(\cdot)$ una densità con media $\mu$ e varianza finita $\sigma^2$. Sia $\overline{X}_n$ la media campionaria di un campione casuale di ampiezza $n$ estratto da $f(\cdot)$. Sia $Z_n$ la variabile casuale definita da

$$Z_n = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}}.$$

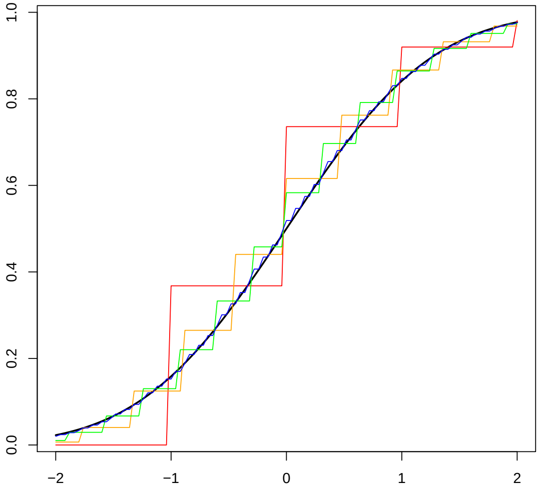
Allora la distribuzione di $Z_n$ si avvicina alla distribuzione normale standardizzata al tendere di $n$ all'infinito.

Verifichiamo il teorema per $f(\cdot) = \frac{e^{-1}(1)^x}{x!} I_{\mathbb{N}}(x)$ (distr. di Poisson con $\lambda = 1$), ricordando che la somma di due variabili i.i.d. di Poisson di parametri $\lambda_1$ e $\lambda_2$ è una Poisson di parametro $\lambda = \lambda_1 + \lambda_2$. Quindi:

$$F_{Z_n}(x) = F_{\mathrm{poisson}(\lambda=n)}(\sqrt{n}x + n).$$

# Il teorema limite centrale (2)

$n = 1, 5, 10, 200$



# Elementi di programmazione

Un semplice ciclo for:

```
for (i in 1:100) {print(i);print(i+1)}
```

Una semplice espressione condizionale:

```
x <- F
if (x == T) print("yes") else print("no")
```

Una semplice funzione:

```
varianza<-function(v) var(v)*(length(v)-1)/length(v)
```

## Problema

Sia
$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2}.$$
Verificare che $f(x)$ è una densità di probabilità.
Definire in R la funzione `rdens(n)` che genera un vettore di $n$ estrazioni campionarie da una variabile aleatoria $X$ con legge $f(x)$.
Confrontare la densità di frequenza di 300 estrazioni casuali da una variabile aleatoria con legge $f(x)$ con la densità di probabilità $f(x)$ (usare `rdens()` e `hist()` oppure `rdens()` e `density()`).