LABORATORIO R

“DISTRIBUZIONI NOTEVOLI”

- Parte 1 di 4 -

DISTRIBUZIONI NOTEVOLI DISCRETE

Nel package standard *stats* sono inclusi i seguenti metodi per le più importanti distribuzioni notevoli discrete.

	Distribuzione	Ripartizione	Quantile	Generazione
Binomiale	dbinom	pbinom	qbinom	rbinom
Ipergeometrica	dhyper	phyper	qhyper	rhyper
Geometrica	dgeom	pgeom	qgeom	rgeom
Poisson	dpois	ppois	qpois	rpois

DISTRIBUZIONI NOTEVOLI DISCRETE

Binomiale

Esercizio 1

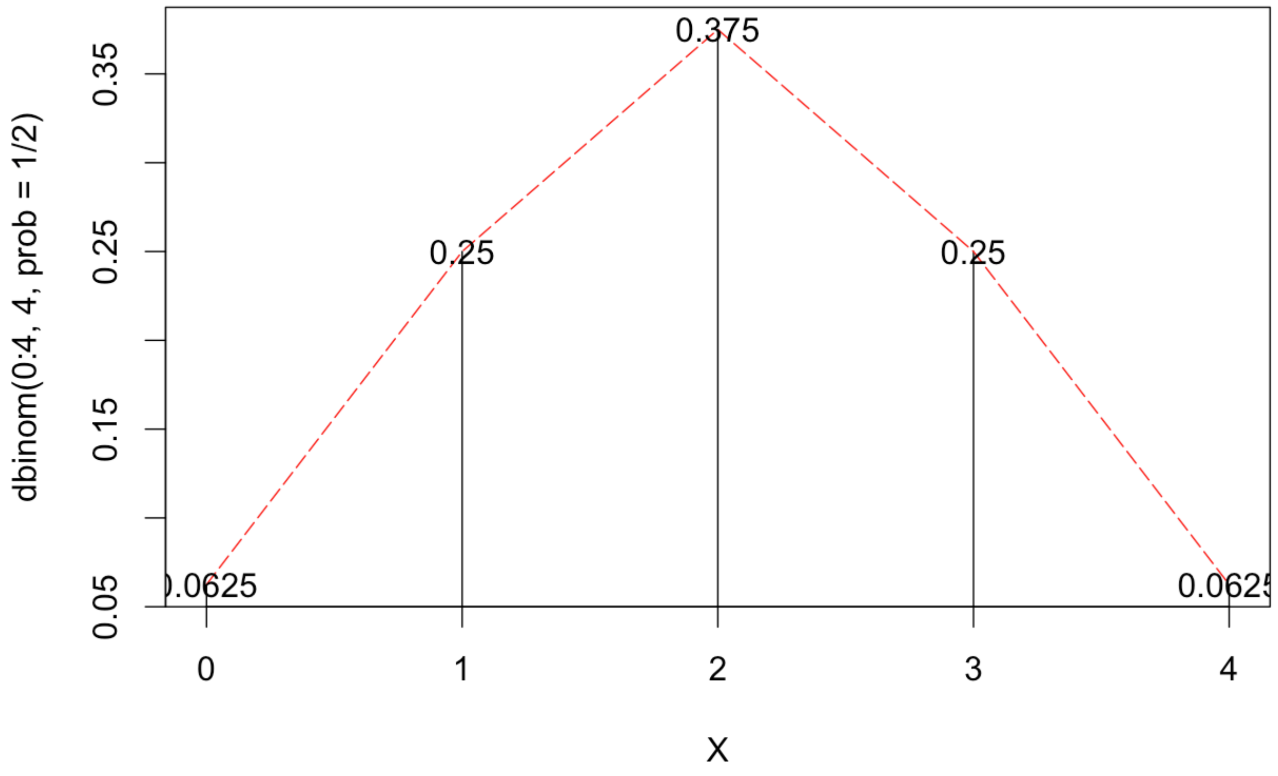
Quattro monete bilanciate vengono lanciate. Assumendo l'indipendenza dei risultati, qual è la probabilità di ottenere due testa e due croce?

```
> dbinom(2,4,prob=1/2)
[1] 0.375  # 1/8
```

DISTRIBUZIONI NOTEVOLI DISCRETE

Disegnare la distribuzione di probabilità della variabile binomiale: X = “numero di volte in cui compare testa”, lanciando 4 volte una moneta bilanciata.

```
> plot(c(0:4),dbinom(0:4,4,prob=1/2),type="h",xlab="X")
> lines(c(0:4),dbinom(0:4,4,prob=0.5) ,lty=5,col="red")
> text(c(0:4), dbinom(0:4,4,prob=0.5),dbinom(0:4,4,prob=0.5))
```



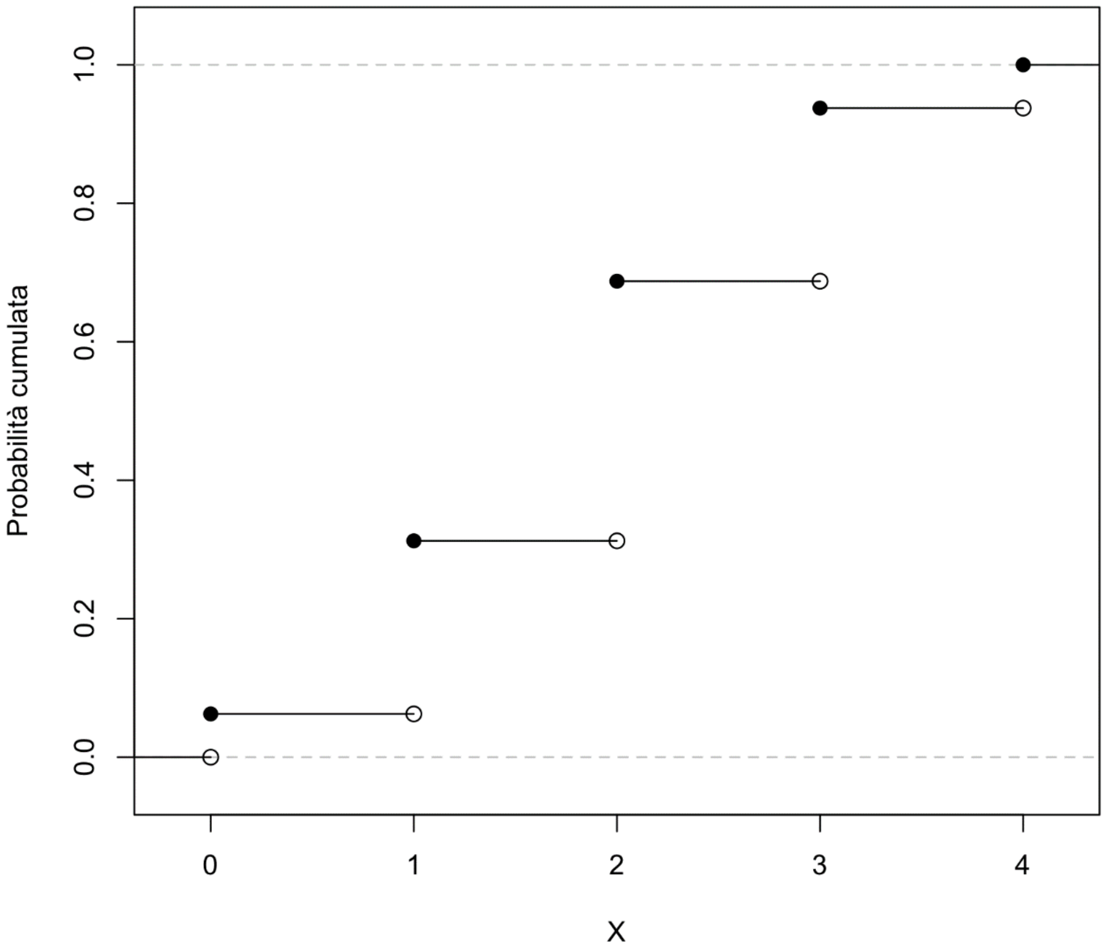
DISTRIBUZIONI NOTEVOLI DISCRETE

Disegnare la funzione di ripartizione

```
# impostare il grafico
> plot(0, xlim = c(-0.2, 4.2), ylim = c(-0.04, 1.04), type = "n", xlab = "X", ylab =
"Probabilità cumulata")

# disegnare due linee orizzontali che limitano la y
> abline(h = c(0,1), lty = 2, col = "grey")

# disegnare una funzione a gradini
> lines(stepfun(0:4, pbinom(-1:4, size = 4, prob = 0.5)), verticals = FALSE, do.p
= FALSE)

# disegnare i punti estremi
> points(0:4, pbinom(0:4, size = 4, prob = 0.5), pch = 16, cex = 1.2)
> points(0:4, pbinom(-1:3, size = 4, prob = 0.5), pch = 1, cex = 1.2)
```

DISTRIBUZIONI NOTEVOLI DISCRETE

Disegnare la funzione di ripartizione



DISTRIBUZIONI NOTEVOLI DISCRETE

Binomiale

Esercizio 2
E’ noto che gli item prodotti da una macchina utensile saranno difettosi con probabilità 0.1. Qual è la probabilità che in un campione di 3 items al più uno sia difettoso?

```
> pbinom(1,3,prob=0.1)
[1] 0.972
```

Esercizio 3
Supponiamo che il colore degli occhi di una persona sia determinato in base ad una coppia di geni e che «d» rappresenti il gene dominante mentre «r» il gene recessivo. Pertanto una persona con la coppia di geni «dd» ha dominanza pura, una con «rr» ha recessione pura e una con «dr» o «rd» è ibrida. Un bambino riceve un gene da ognuno dei genitori. Se rispetto al colore degli occhi i due genitori «ibridi» hanno 4 figli.
Qual è la probabilità che esattamente 3 dei 4 figli abbiano almeno 1 gene dominante?

```
> dbinom(3,4,prob=3/4)
[1] 0.421875
```

DISTRIBUZIONI NOTEVOLI DISCRETE

Binomiale

Esercizio 6
Un'azienda produce dischetti, la probabilità che un dischetto sia difettoso è pari a 0.01. L'azienda vende i dischetti in confezioni da 10 e rimborsa l'acquirente se più di 1 dischetto è difettoso.

1. Quale proporzione delle confezioni sarà restituita?

```
> 1 - pbinom(1, 10, 0.01)
[1] 0.0042662
# Oppure
> pbinom(1, 10, 0.01, lower.t=FALSE)
```

2. Se un acquirente acquista 3 scatole, qual è la probabilità che ne restituisca esattamente 1?

```
> dbinom(1, 3, 0.0042662)
[1] 0.01268963
```

DISTRIBUZIONI NOTEVOLI DISCRETE

Poisson

Esercizio 4
Supponiamo che il numero di errori tipografici presenti in una singola pagina di un libro sia distribuito secondo una Poisson con parametri λ = 1.
Si calcoli la probabilità che vi sia almeno un errore in una data pagina.

```
> ppois(0,lambda=1,lower.t=FALSE)
[1] 0.6321206
# Oppure
> 1- ppois(0,1)
```

DISTRIBUZIONI NOTEVOLI DISCRETE

Poisson

Esercizio 7
Si supponga che la probabilità che un prodotto costruito da una macchina sia difettoso è pari a 0.1.
Si trovi la probabilità che un campione di 10 prodotti contenga al più un prodotto difettoso.

```
> ppois(1,lambda=10*0.1)
[1] 0.7357589
```

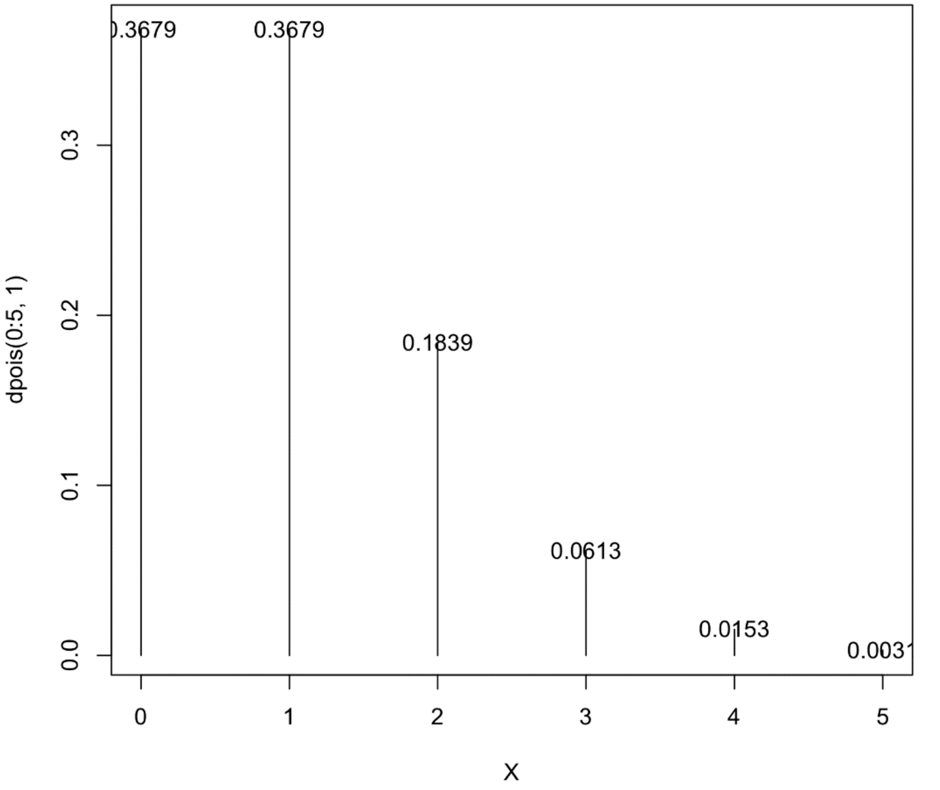
Si trovi la probabilità che un campione di 10 prodotti contenga tra 3 e i 5 prodotti difettosi

```
> diff(ppois(c(2, 5), lambda = 10*0.1))
[1] 0.07970721
#Equivale a:
> dpois(3,lambda= 10*0.1)+dpois(4,lambda= 10*0.1)+dpois(5,lambda= 10*0.1)
```

DISTRIBUZIONI NOTEVOLI DISCRETE

Tracciare il grafico della distribuzione della variabile casuale di Poisson per $\lambda=1$

```
> plot(c(0:5),dpois(0:5,1),type="h",xlab="X")
> text(c(0:5), dpois(0:5,1), round(dpois(0:5,1),4))
```



DISTRIBUZIONI NOTEVOLI DISCRETE

Geometrica

Esempio
Un calciatore ha una probabilità di segnare un calcio di rigore di 0.812.
Si supponga che tirare un calcio di rigore sia un evento Bernoulliano.

Qual è la probabilità che sbagli 5 calci di rigore prima di segnarne uno.

```
> dgeom(5, prob = 0.812)
[1] 0.0001906976
```

Qual è la probabilità che sbagli al massimo 5 calci di rigore prima di segnarne uno.

```
> pgeom(5, prob = 0.812)
[1] 0.9999558
```

Qual è la probabilità che sbagli almeno 5 calci di rigore prima di segnarne uno.

```
> pgeom(4, prob = 0.812, lower.tail = FALSE)
[1] 0.0002348493
```

LABORATORIO R

"DISTRIBUZIONI NOTEVOLI"

- Parte 2 di 4 -

DISTRIBUZIONI NOTEVOLI CONTINUE (parte 1)

Nel package standard stats sono inclusi i seguenti metodi per le più importanti distribuzioni notevoli continue.

	Densità	Ripartizione	Quantile	Generazione
Uniforme	dunif	punif	qunif	runif
Esponenziale	dexp	pexp	qexp	rexp

DISTRIBUZIONI NOTEVOLI CONTINUE

Il package distr contiene classi per gestire una grande varietà di distribuzioni:

Distribution
UnivariateDistribution
UnivarMixingDistribution
UnivarLebDecDistribution
AffLinUnivarLebDecDistribution
CompoundDistribution
AbscontDistribution
AffLinAbscontDistribution
Arcsine
Beta
Cauchy
ExpOrGammaOrChisq
Exp
Gammad
Chisq
Fd
Lnorm
Logis
Norm
Td
Unif
Weibull
DiscreteDistribution
AffLinDiscreteDistribution
LatticeDistribution
AffLinLatticeDistribution
Binom
Dirac
Hyper
NBinom
Geom
Pois

DISTRIBUZIONI NOTEVOLI CONTINUE

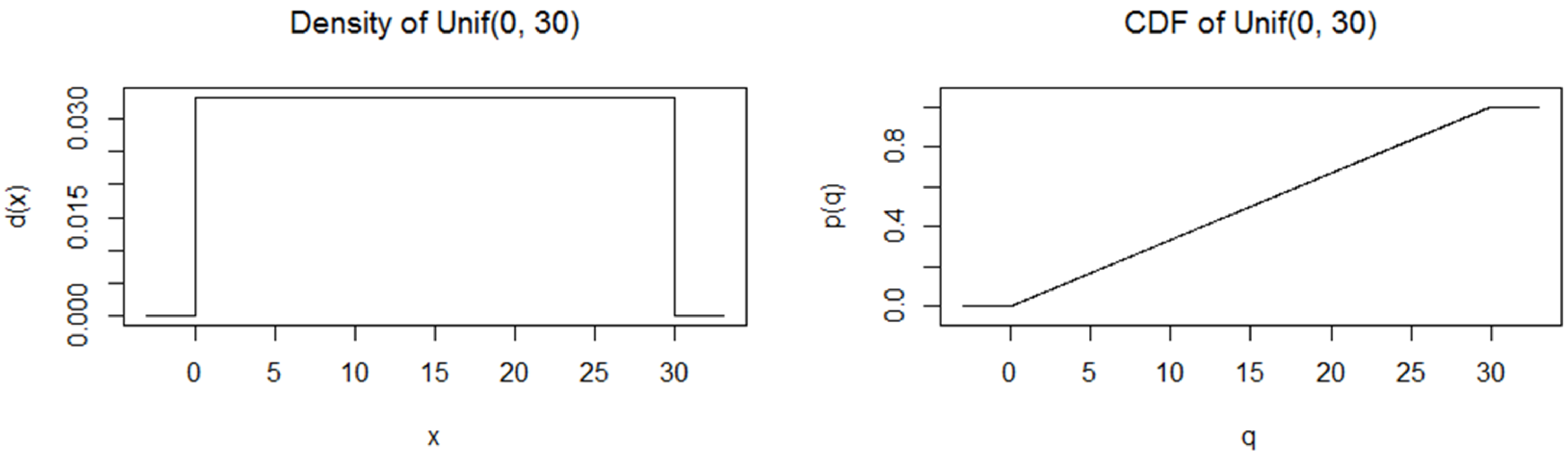
Uniforme

Esempio 1
Sia X una v.c. che rappresenta la probabilità di ricevere una telefonata tra le 10 e le 10:30. Si calcoli la probabilità che la chiamata arrivi tra le 10:10 e le 10:20.
X è distribuita secondo una variabile uniforme con parametri a = 0 e b = 30.

```
> diff(punif(c(10,20),0,30))
>[1] 0.3333333

# oppure
> punif(20,0,30) - punif(10,0,30)
```

DISTRIBUZIONI NOTEVOLI CONTINUE

```
> library(distr)
> X <- Unif(Min = 0, Max = 30)
> plot(X, to.draw.arg=c("d","p"))
```



DISTRIBUZIONI NOTEVOLI CONTINUE

Generare 10 numeri da una distribuzione uniforme compresa tra 0 e 50

```
> runif(n = 10, min = 0, max = 50)
 [1] 43.258352  8.419059  6.827797 23.992698 19.979657 30.773017  4.216128
 [8] 48.204129 40.322714 25.301487
```

DISTRIBUZIONI NOTEVOLI CONTINUE

Esercizio 08

Un bus arriva ad una data fermata ad intervalli di 15 minuti a partire dalle ore 7:00. Poiché il bus passa ogni quarto d'ora, se un passeggero arriva alla fermata in un istante di tempo uniformemente distribuito nell'intervallo 7:00-7:30, si determinino:

a) Probabilità che attenda meno di 5 minuti
b) Probabilità che attenda almeno 12 minuti

a)

```
> (punif(15,0,30) - punif(10,0,30)) + (punif(30,0,30) – punif(25,0,30))
[1] 0.3333333
```

b)

```
> (punif(3,0,30) - punif(0,0,30)) + (punif(18,0,30) - punif(15,0,30))
[1] 0.2
```

DISTRIBUZIONI NOTEVOLI CONTINUE

Esponenziale

Esempio 2
Sia X una v.c. che rappresenta la durata di una batteria in mesi. X è distribuita secondo una esponenziale con parametro $\boldsymbol{\lambda = 0.2}$.

Si calcoli la probabilità che una batteria duri tra i 4 e i 6 mesi.

```
> diff(pexp(c(4,6),rate = 0.2))
[1] 0.1481348
```

Si calcoli la probabilità che una batteria duri almeno 10 mesi.

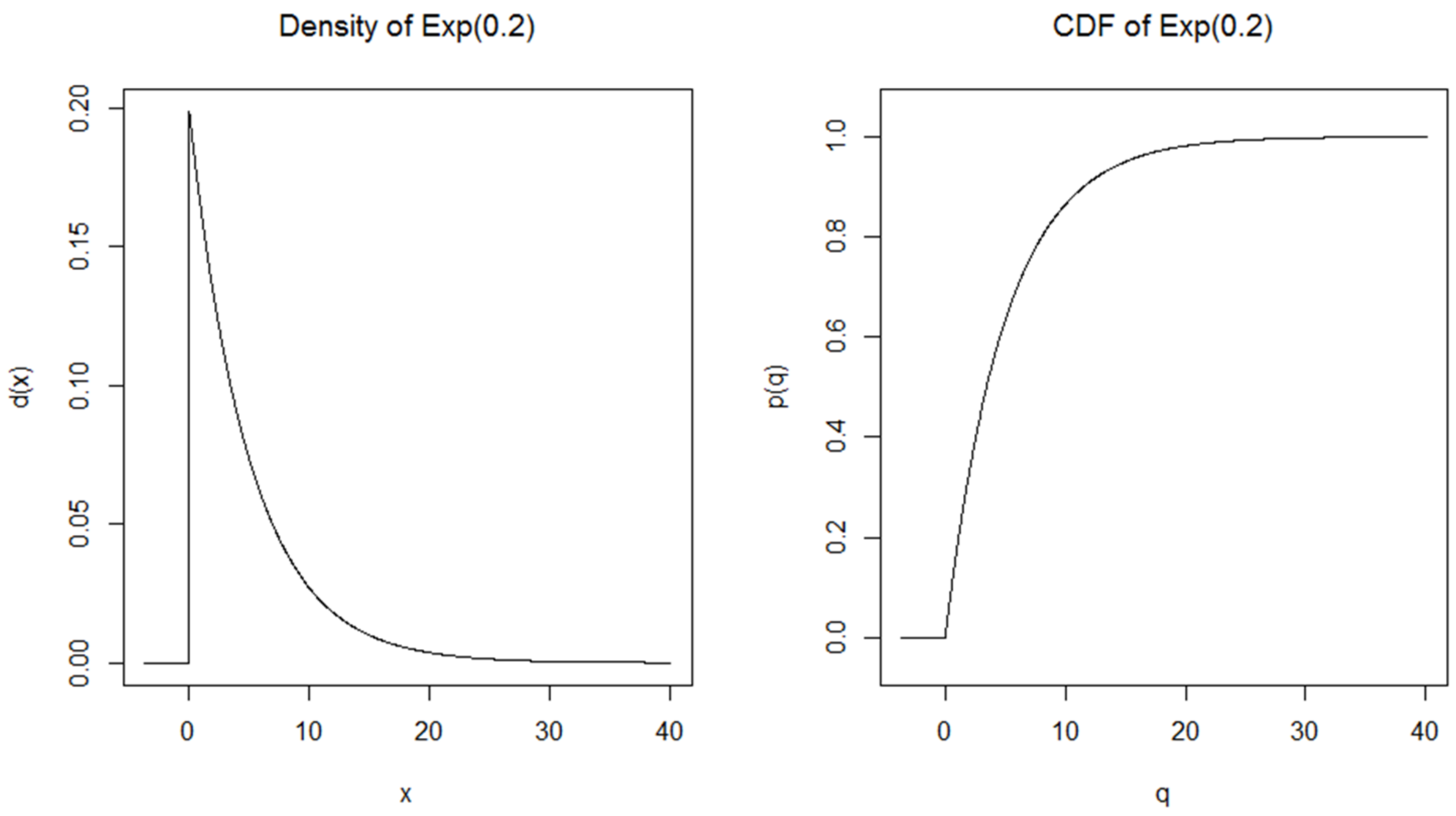
```
> pexp(10,rate = 0.2,lower.tail = FALSE)
[1] 0.1353353
```

Si calcoli la probabilità che una batteria duri altri 10 mesi sapendo che la batteria sta durando da 4 mesi.

```
> pexp(14,rate = 0.2,lower.tail = FALSE)/pexp(4,rate = 0.2,lower.tail = FALSE)
```

DISTRIBUZIONI NOTEVOLI CONTINUE

```
> plot(Exp(rate = 0.2), to.draw.arg = c("d","p"))
```



DISTRIBUZIONI NOTEVOLI CONTINUE

Generare 10 numeri da una distribuzione esponenziale con $\lambda=0.4$

```
> rexp(10, rate=0.4)
 [1] 0.88158461 3.80154731 3.93130611
 [4] 1.11978673 2.39023489 2.91561690
 [7] 0.01601692 5.59117098 1.61702767
[10] 2.41262227
```

LABORATORIO R

"DISTRIBUZIONI NOTEVOLI"

- Parte 3 di 4 -

DISTRIBUZIONI NOTEVOLI CONTINUE (parte 2)

Nel package standard stats sono inclusi i seguenti metodi per le più importanti distribuzioni notevoli continue.

	Densità	Ripartizione	Quantile	Generazione
Normale	dnorm	pnorm	qnorm	rnorm

DISTRIBUZIONI NOTEVOLI CONTINUE

Normale

Esempio 3.1 (teoria)

Sia X una variabile aleatoria con distribuzione normale di parametri μ=10 e σ=1 e si voglia determinare la probabilità dell'evento

$$"X \in [9.2, 11.35]"$$

```
> pnorm(11.35,mean=10) - pnorm(9.2,mean=10)
[1] 0.6996366
```

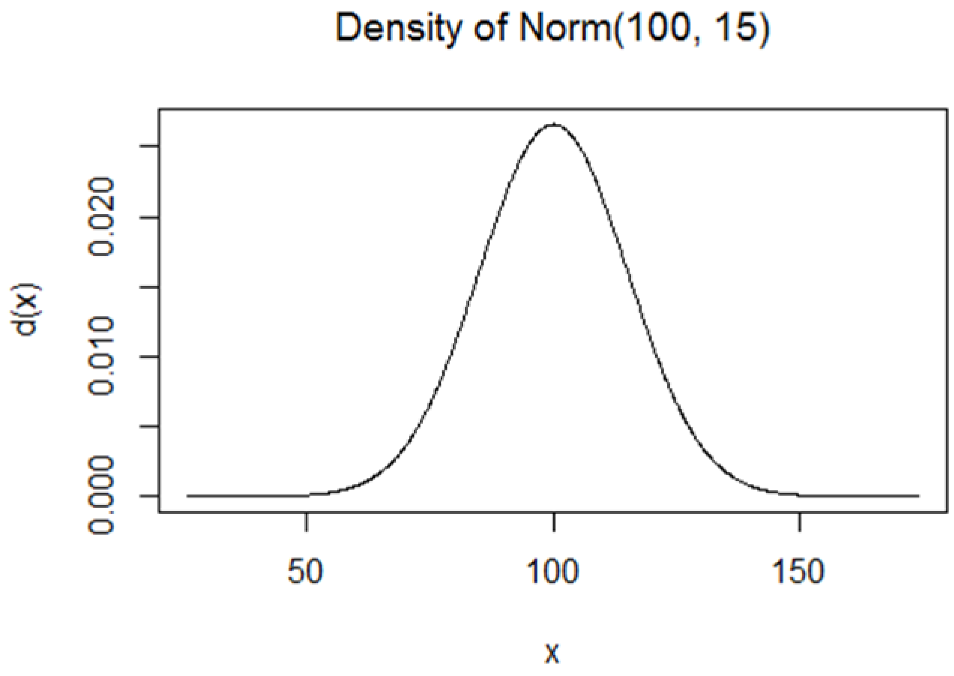
DISTRIBUZIONI NOTEVOLI CONTINUE

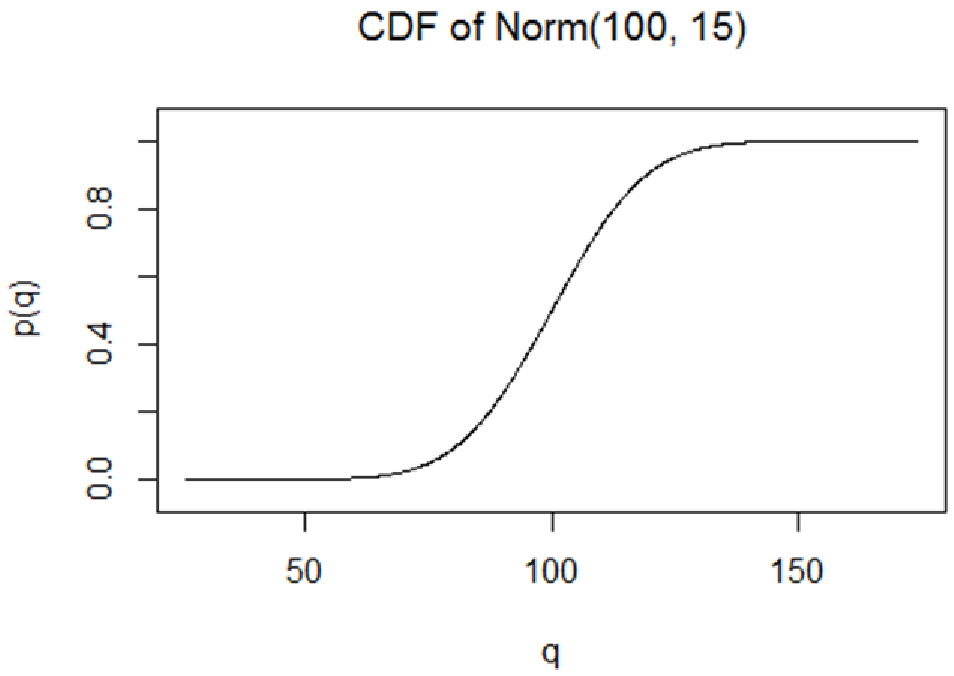
\# Esempio
Sia X una v.c. che rappresenta il risultato ad un test di QI. X è distribuita secondo una variabile normale con media 100 e deviazione standard 15.
Si calcoli la probabilità che una persona abbia un QI tra 85 e 115.

```
> diff(pnorm(c(85,115),mean = 100,sd = 15))
[1] 0.6826895
```

DISTRIBUZIONI NOTEVOLI CONTINUE

```
> library(distr)
> X <- Norm(mean = 100, sd = 15)
> plot(X, to.draw.arg=c("d","p"))
```





DISTRIBUZIONI NOTEVOLI CONTINUE

$$\mathrm{P}\left(X \in [\mu - \sigma, \mu + \sigma]\right) = 0.683$$

$$\mathrm{P}\left(X \in [\mu - 2 \cdot \sigma, \mu + 2 \cdot \sigma]\right) = 0.954$$

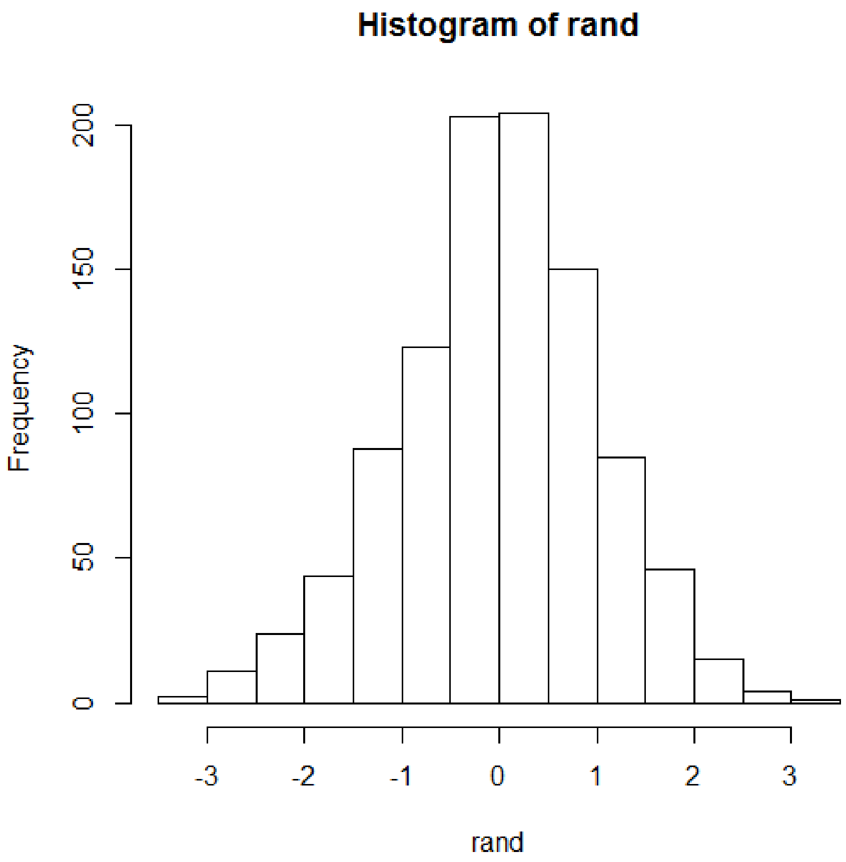
$$\mathrm{P}\left(X \in [\mu - 3 \cdot \sigma, \mu + 3 \cdot \sigma]\right) = 0.997$$

```
> pnorm(1:3) - pnorm(-(1:3))
[1] 0.6826895 0.9544997 0.9973002
```

DISTRIBUZIONI NOTEVOLI CONTINUE

Generare 1000 numeri da una normale standardizzata

```
> rand <- rnorm(1000)
> hist(rand)
```



DISTRIBUZIONI NOTEVOLI CONTINUE

```
# Esempio precedente (quantili)
```

Qual è il minimo QI che una persona deve avere per essere nell'1% delle persone con il QI più alto?

```
> qnorm(0.99, mean = 100, sd = 15)
[1] 134.8952
```

DISTRIBUZIONI NOTEVOLI CONTINUE

Calcolare $z_{0.025}$,$z_{0.01}$ e $z_{0.05}$

```
> qnorm(c(0.025, 0.01, 0.005),lower.tail = FALSE)
[1] 1.959964 2.326348 2.575829

#oppure
> qnorm(c(0.975,0.99,0.995))
```

DISTRIBUZIONI NOTEVOLI CONTINUE

\# Esercizio 09

Se X è una variabile distribuita secondo una normale con $\mu = 3$ e $\sigma^2 = 16$, si determini:

a) P(X < 11)
b) P(X > -1)
c) P(2 < X < 7)

a)
```
> pnorm(11, mean=3, sd=sqrt(16))
[1] 0.9772499
```
b)
```
> pnorm(-1, mean=3, sd=4, lower.tail=FALSE)
[1] 0.8413447
> 1 - pnorm(-1, mean=3, sd=4) #oppure
```
c)
```
> pnorm(7, mean=3,sd=4) - pnorm(2, mean=3,sd=4)
[1] 0.4400511
```

LABORATORIO R

"DISTRIBUZIONI NOTEVOLI"

- Parte 4 di 4 -

DISTRIBUZIONI NOTEVOLI CONTINUE (parte 3)

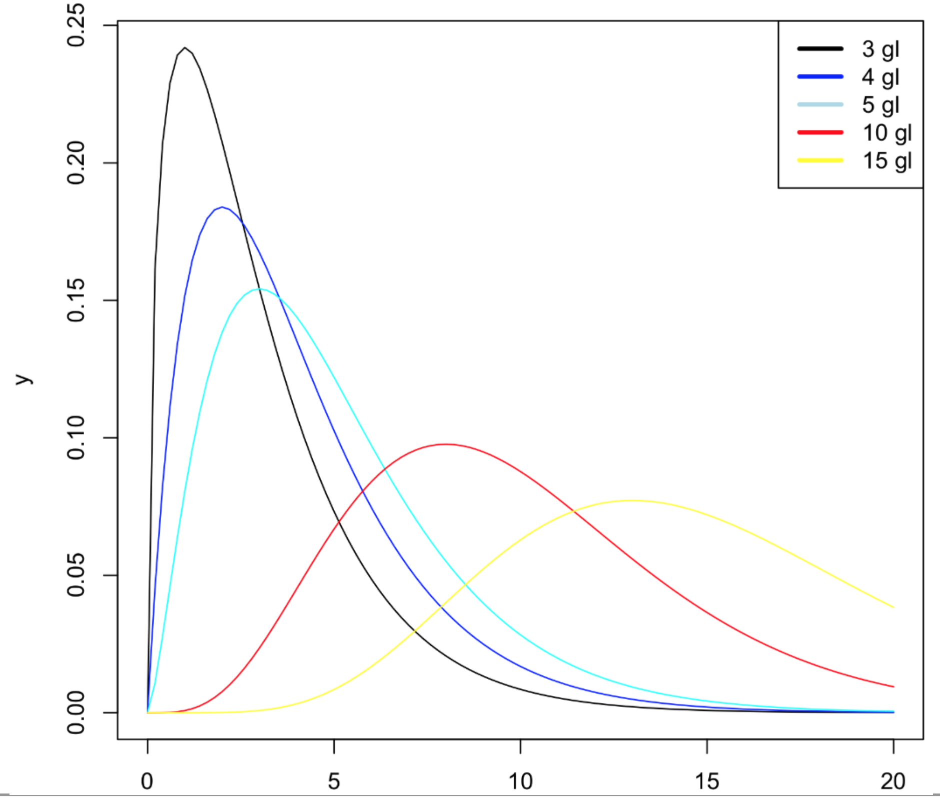
Nel package standard stats sono inclusi i seguenti metodi per le più importanti distribuzioni notevoli continue.

	Densità	Ripartizione	Quantile	Generazione
Chi-Quadro	dchisq	pchisq	qchisq	rchisq
T di Student	dt	pt	qt	rt
F di Fisher	df	pf	qf	rf

DISTRIBUZIONI NOTEVOLI CONTINUE

Chi-Quadro

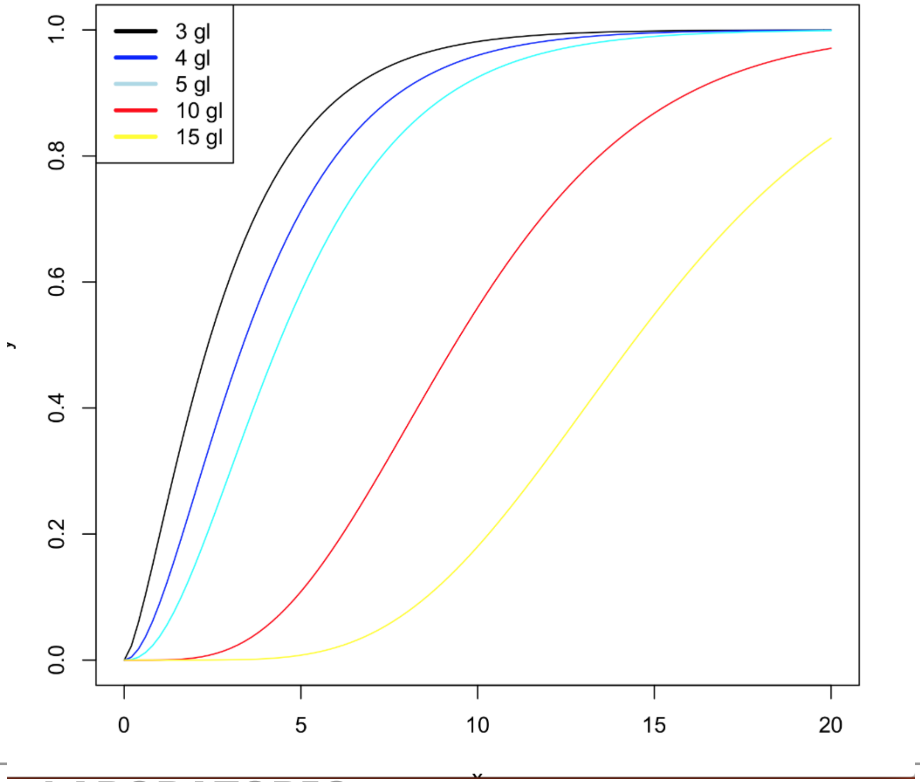
```
> curve(dchisq(x, df = 3), from = 0, to = 20, ylab = "y")
> ind <- c(4, 5, 10, 15)
> for (i in ind) curve(dchisq(x, df = i), 0, 20, add = TRUE,col=i)
> legend("topleft",legend=c("3 gl","4 gl","5 gl","10 gl","15 gl"),
+       col=c("Black","Blue","light Blue","Red","Yellow"), lty=c(1,1,1,1,1), lwd=c(3,3,3,3,3))
```



DISTRIBUZIONI NOTEVOLI CONTINUE

Chi-Quadro

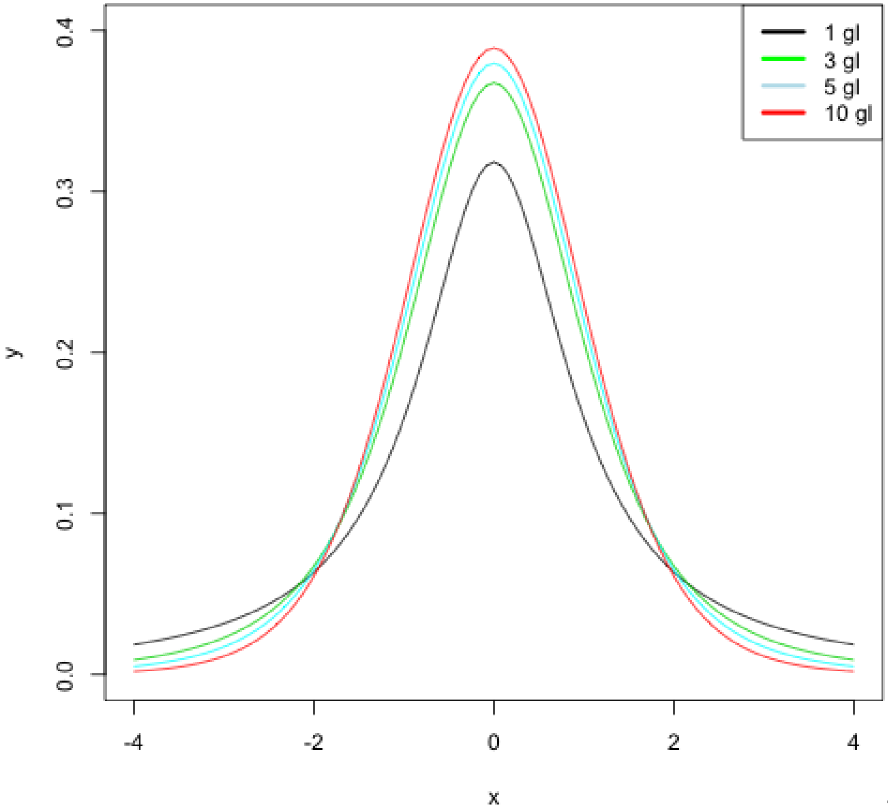
```
> curve(pchisq(x, df = 3), from = 0, to = 20, ylab = "y")
> ind <- c(4, 5, 10, 15)
> for (i in ind) curve(pchisq(x, df = i), 0, 20, add = TRUE, col=i)
> legend("topleft",legend=c("3 gl","4 gl","5 gl","10 gl","15 gl"),
+      col=c("Black","Blue","light Blue","Red","Yellow"), lty=c(1,1,1,1,1), lwd=c(3,3,3,3,3))
```



DISTRIBUZIONI NOTEVOLI CONTINUE

T di Student

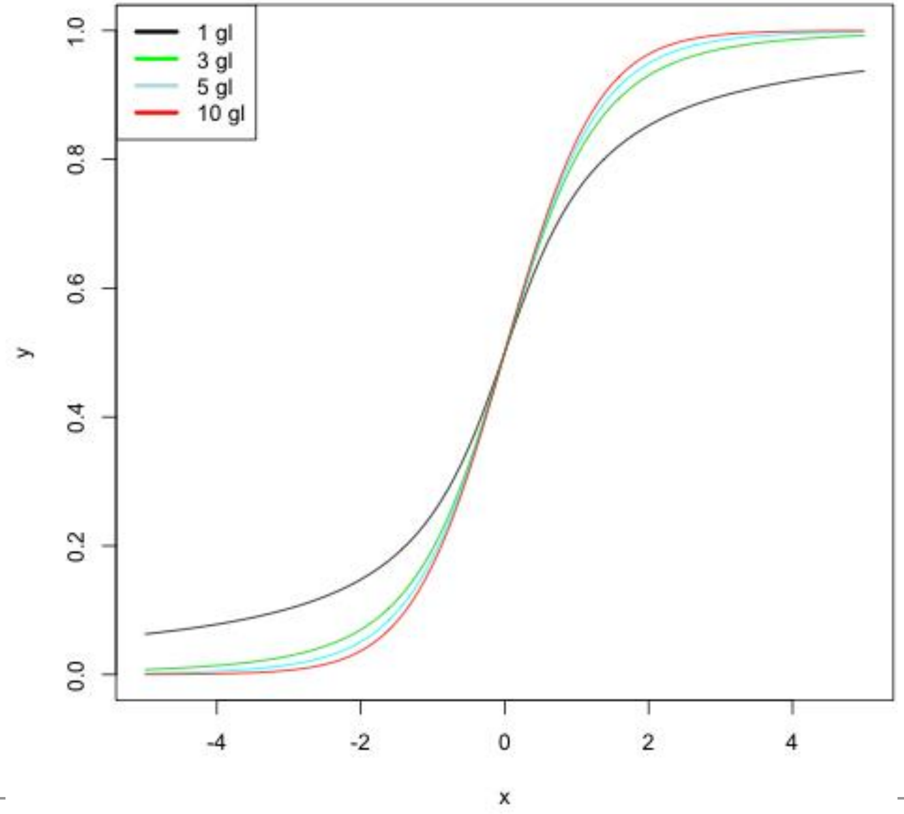
```
> curve(dt(x, df = 1), from = -4, to = 4, ylim = c(0,0.4), ylab = "y")
> ind <- c(3, 5, 10)
> for (i in ind) curve(dt(x, df = i), -4, 4, add = TRUE,col=i)
> legend("topright",legend=c("1 gl","3 gl","5 gl","10 gl"),
+      col=c("Black","Green","light Blue","Red"),lty=c(1,1,1,1), lwd=c(3,3,3,3))
```



DISTRIBUZIONI NOTEVOLI CONTINUE

T di Student

```
> curve(pt(x, df = 1), from = -4, to = 4, ylim = c(0,1), ylab = "y")
> ind <- c(3, 5, 10)
> for (i in ind) curve(dt(x, df = i), -4, 4, add = TRUE,col=i)
> legend("topright",legend=c("1 gl","3 gl","5 gl","10 gl"),
+       col=c("Black","Green","light Blue","Red"),lty=c(1,1,1,1), lwd=c(3,3,3,3))
```

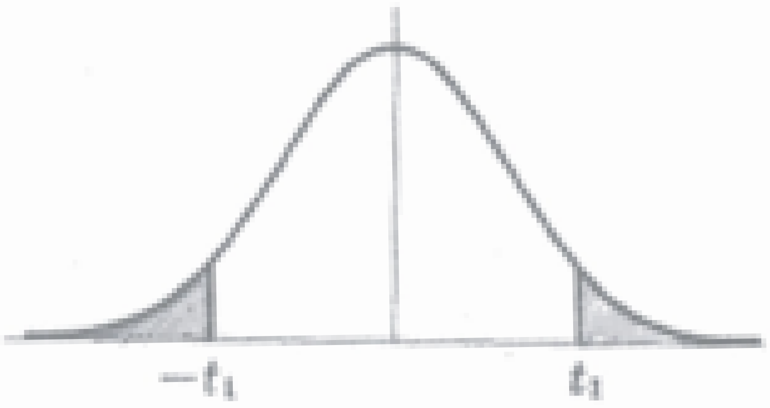


DISTRIBUZIONI NOTEVOLI CONTINUE

\# Esempio
Il grafico della distribuzione t di Student con 9 gradi di libertà è riportato in figura.Trovare il valore di t_1 tale per cui:

a) l'area a destra è pari a 0.05
b) il totale dell'area ombreggiata è pari a 0.05
c) il totale dell'area non ombreggiata è pari a 0.99

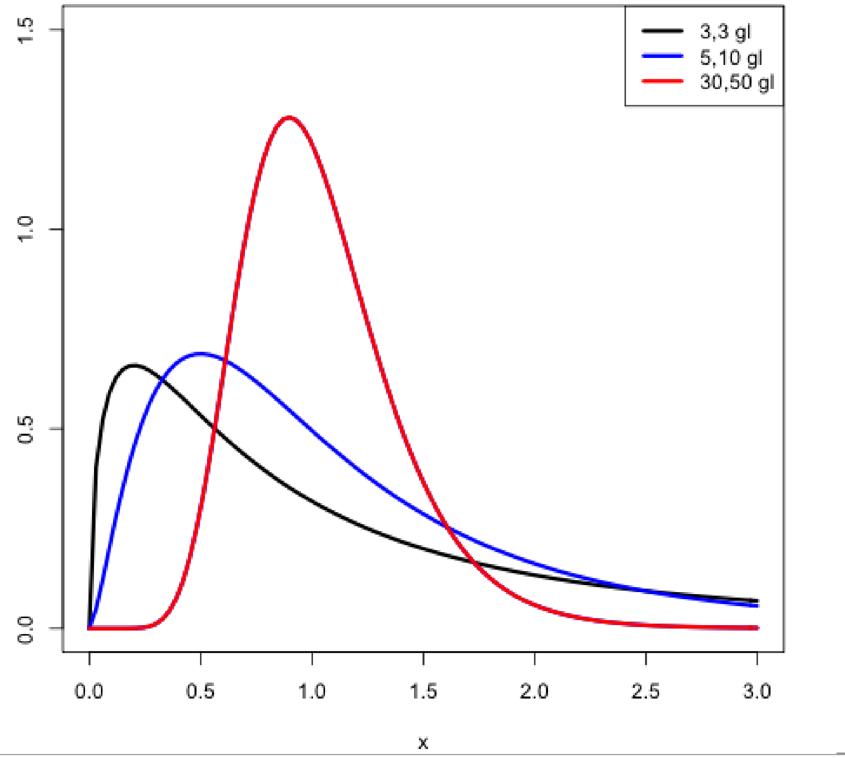
a)
```
> qt(0.95,df=9)
[1] 1.833113
```
b)
```
> qt(0.975, df =9)
[1] 2.262157
```
c)
```
> qt(0.995, df =9)
[1] 3.249836
```



DISTRIBUZIONI NOTEVOLI CONTINUE

F di Fisher

```
> curve(df(x,df1=3,df2=3),0,+3,col="Black",lwd=3)
> curve(df(x,df1=5,df2=10),0,+3,col="Blue",lwd=3,add=TRUE)
> curve(df(x,df1=30,df2=50),0,+3,col="Red",lwd=3,add=TRUE)
> legend("topright",legend=c("3,3 gl","5,10 gl","30,50 gl"),
    col=c("Black","Blue","Red"),lty=c(1,1,1), lwd=c(3,3,3))
```



DISTRIBUZIONI NOTEVOLI CONTINUE

F di Fisher

```
> curve(pf(x,df1=3,df2=3),0,10,col="Black",lwd=3,ylim=c(0,1))
> curve(pf(x,df1=5,df2=10),0,10,col="Blue",lwd=3,add=TRUE)
> curve(pf(x,df1=30,df2=50),0,10,col="Red",lwd=3,add=TRUE)
> legend("topleft",legend=c("3,3 gl","5,10 gl","30,50 gl"),
   col=c("Black","Blue","Red"),lty=c(1,1,1), lwd=c(3,3,3))
```