Parte 6

Modelli di probabilità per variabili casuali continue

Variabile casuale **Uniforme continua** (o Rettangolare)

Esperimento casuale che *genera* un numero reale compreso in un *intervallo limitato*, in modo tale che (per ragioni di simmetria e indifferenza) **sotto-intervalli** di *pari ampiezza* abbiano la **medesima probabilità**

Implicazione	Probabilità di un qualsiasi intervallo reale **proporzionale** alla *ampiezza* dell'intervallo medesimo

Ambiti applicativi:

Situazioni artificiali come giochi ed esperimenti geometrici, *teoria della simulazione*

Variabile casuale continua *associata* all'esperimento casuale	$X =$ "**Numero** generato dall'esperimento casuale"

$x \in [a, b]$	$a, b \in \mathbb{R}$ (numeri reali qualsiasi)	$[a, b]$ **supporto** della v.c.

a e b sono i **parametri** che *caratterizzano* tale v.c.

L'**insieme** delle v.c. *generato* da *tutti i possibili* valori di a e b (sono una *infinità non numerabile*) è la **famiglia parametrica** di v.c.

Funzione/Modello di **densità** parametrica/o della v.c. X			
$f(x;a,b)=\dfrac{1}{b-a}$	$x\in[a,b]$	$f(x;a,b)=0$ altrove	$(a,b)\in\mathbb{R}^2$ $\mathbb{R}^2$ spazio parametrico
$X\sim U(a,b)$			
La v.c. X ha distribuzione di probabilità **Uniforme continua** di parametri a e b			

Proprietà soddisfatte dalla funzione di densità	
$f(x;a,b)=\dfrac{1}{b-a}>0$	$\int_a^b f(x;a,b)\,dx=\int_a^b \dfrac{1}{b-a}dx=\dfrac{b-a}{b-a}=1$

L'**insieme** delle $f(x;a,b)$ *generato* dai valori di a e b è la **famiglia parametrica** di funzioni di **densità** (modelli di probabilità) di tipo **Uniforme continuo** di parametri a e b

Funzione di ripartizione *parametrica* di una v.c. $X \sim U(a, b)$

$$F(x) = F(x; a, b) = \int_a^x f(t; a, b)\, dt = \int_a^x \frac{1}{b-a}\, dt = \frac{x-a}{b-a} = -\frac{a}{b-a} + \frac{1}{b-a} x$$

$x \in [a, b]$	Funzione **lineare crescente** su $[a, b]$	In generale è una funzione **non decrescente**

$F(x; a, b) = 0$	se $x \leq a$	$F(x; a, b) = 1$	se $x \geq b$

Siano $x_1, x_2 \in [a, b]$ due generici valori del supporto di X tali che $x_1 < x_2$

$$P(x_1 \leq X \leq x_2) = F(x_2; a, b) - F(x_1; a, b) = \frac{x_2 - a}{b-a} - \frac{x_1 - a}{b-a} = \frac{x_2 - x_1}{b-a}$$

La probabilità dell'intervallo reale $[x_1, x_2]$ dipende solo (in modo **proporzionale**) dalla *ampiezza* $x_2 - x_1$ dell'intervallo	Non dipende dagli specifici valori x_1 e x_2
Il *fattore di proporzionalità* è la funzione di densità	$f(x; a, b) = \dfrac{1}{b-a}$

Intervalli di *uguale ampiezza* contenuti nel supporto $[a, b]$ hanno *uguale probabilità*

Media e **varianza** di una v.c. $X \sim U(a,b)$

$$f(x;a,b) = \frac{1}{b-a} \qquad x \in [a,b] \qquad (a,b) \in \mathbb{R}^2$$

$$E(X) = \int_a^b x \frac{1}{b-a} dx = \frac{1}{b-a} \int_a^b x\, dx = \frac{1}{b-a} \left[\frac{x^2}{2} \right]_a^b = \frac{1}{b-a} \frac{b^2-a^2}{2} = \frac{a+b}{2}$$

La **media** coincide con il **valore centrale** del supporto

$$V(X) = \frac{(b-a)^2}{12}$$

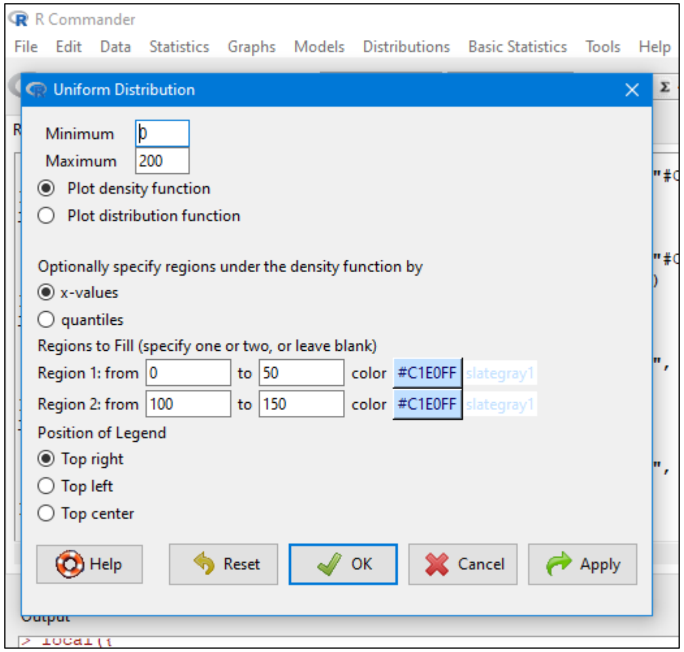
Media e varianza *funzioni* di entrambi i parametri a e b

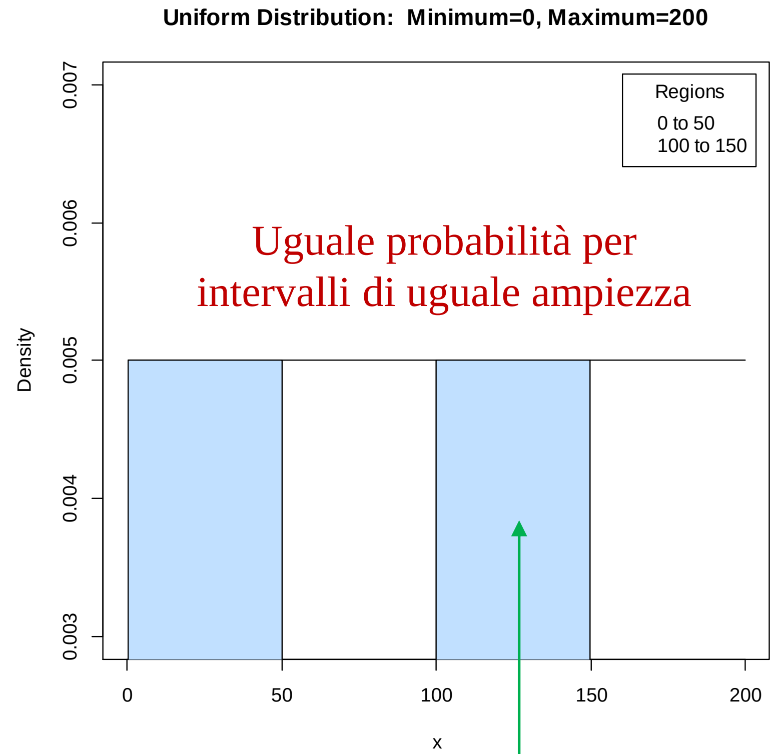
Esempio 11	**Esperimento casuale**
Lancio (con forza) di una palla su un tavolo da biliardo lungo 200 cm. La palla rimbalzerà tra le varie sponde sino a fermarsi.	

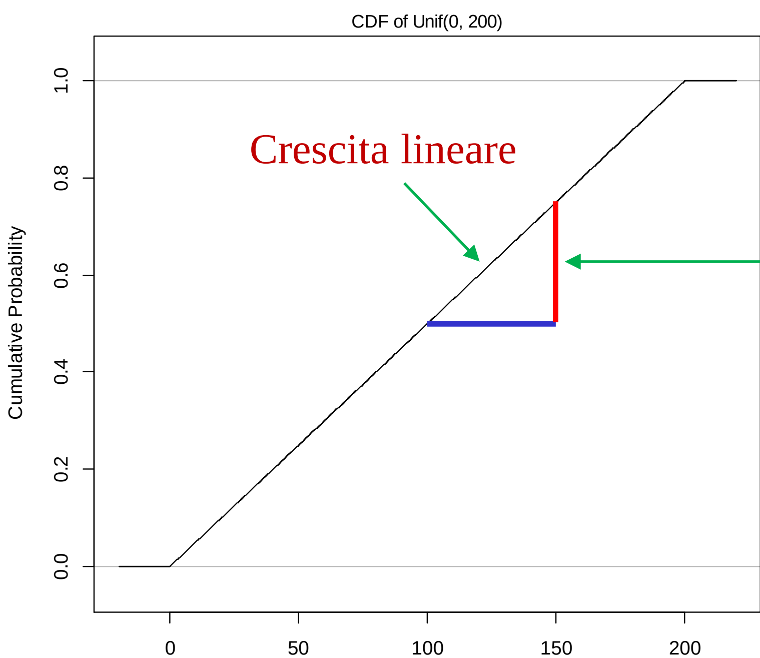
$X =$ "**Distanza** (minima) tra il **bordo** del tavolo e il **punto** in cui la palla si è fermata"

X è una v.c. **Uniforme continua** in $[0{,}200]$		Parametri $a = 0, b = 200$	
$X \sim U(0{,}200)$	$f(x) = \dfrac{1}{200}$	$x \in [0{,}200]$	$f(x) = 0$ altrove
$F(x) = \dfrac{x}{200}$	$x \in [0{,}200]$	$E(X) = 100$	
$F(x) = 0$ $x \leq 0$	$F(x) = 1$ $x \geq 200$	$V(X) = \dfrac{200^2}{12} = 3333.\overline{3}$	

Distribuzione simmetrica $Me = E(X) = 100$







Nella *teoria della simulazione*,
ha un ruolo rilevante la v.c. $U(0,1)$

Variabile casuale **Normale** (o **di Gauss** o **Gaussiana**)

Variabile casuale continua che *può assumere* come valore un **qualsiasi** numero reale (tra $-\infty$ e ∞)

Il **supporto** di tale v.c. è l'intero **asse reale** $\mathbb{R}$

Introdotta nella storia della probabilità come distribuzione degli **errori accidentali**

Carl Friedrich **Gauss**
1777 – 1855

Altri contributi:
Galileo (1632), De Moivre (1733), Daniel Bernoulli (1770), Laplace (1810)

La presenza della v.c. Normale nelle applicazioni e nella teoria, come *suggerisce* la sua denominazione, è quasi una *regola* (la "norma"!)

Moltissime v.c. continue (interpreti di fenomeni reali) con supporto **reale** (ma anche con supporto **limitato** o **non negativo**, ndr) sono *ben definite* come v.c. Normali

In questo caso, naturalmente, l'**esperimento casuale** di riferimento, cui è *associata* la v.c. Normale, *genera* un numero reale *qualsiasi* (o *limitato* o *non negativo*)

Variabile casuale Normale *estremamente frequente* come v.c. di **popolazione**

Importanza della v.c. **Normale** negli **sviluppi teorici** dei **metodi inferenziali**

Variabili casuali **campionarie** *connesse* alla v.c. Normale

Variabili casuali continue *generate* dal campionamento casuale, e **derivate** dalla v.c. Normale, con distribuzione di probabilità **nota**

Importanza cruciale per l'applicazione *dei più comuni* metodi inferenziali

Variabili casuali **Chi-quadrato**, t **di Student**, F **di Fisher**

Variabili casuali **campionarie** *asintoticamente* Normali

Variabili casuali *generate* dal campionamento casuale, con distribuzione di probabilità **approssimativamente** Normale se l'ampiezza campionaria n è *sufficientemente grande* (se si dispone di un **grande campione**)

Motivazione	Risultati **limite** sulla loro distribuzione di probabilità (che valgono cioè per $n \to \infty$)

Importanza cruciale per l'applicazione *generale* dei metodi inferenziali

Risultato **limite** principale	**Teorema del limite centrale**

X v.c. Normale	$x \in \mathbb{R} = (-\infty, \infty)$	$\mathbb{R}$ **supporto** della v.c.

Funzione/Modello di **densità** parametrica/o della v.c. X

Notazione $exp\{b\} = e^b$	$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right\}$	$X \sim N(\mu, \sigma^2)$

La v.c. X ha distribuzione di probabilità **Normale** di parametri μ e σ^2

L'**insieme** delle $f(x; \mu, \sigma^2)$ *generato* dai valori di μ e σ^2 è la **famiglia parametrica** di funzioni di **densità** (modelli di probabilità) di tipo **Normale** di parametri μ e σ^2

L'**insieme** corrispondente delle v.c. è la **famiglia parametrica** di v.c. **Normali**

I parametri μ e σ^2 **coincidono**, rispettivamente, con la **media** e la **varianza** della v.c. X

$E(X) = \mu$	$\mu \in \mathbb{R} = (-\infty, \infty)$	$V(X) = \sigma^2$	$\sigma^2 \in \mathbb{R}_0^+ = [0, \infty)$

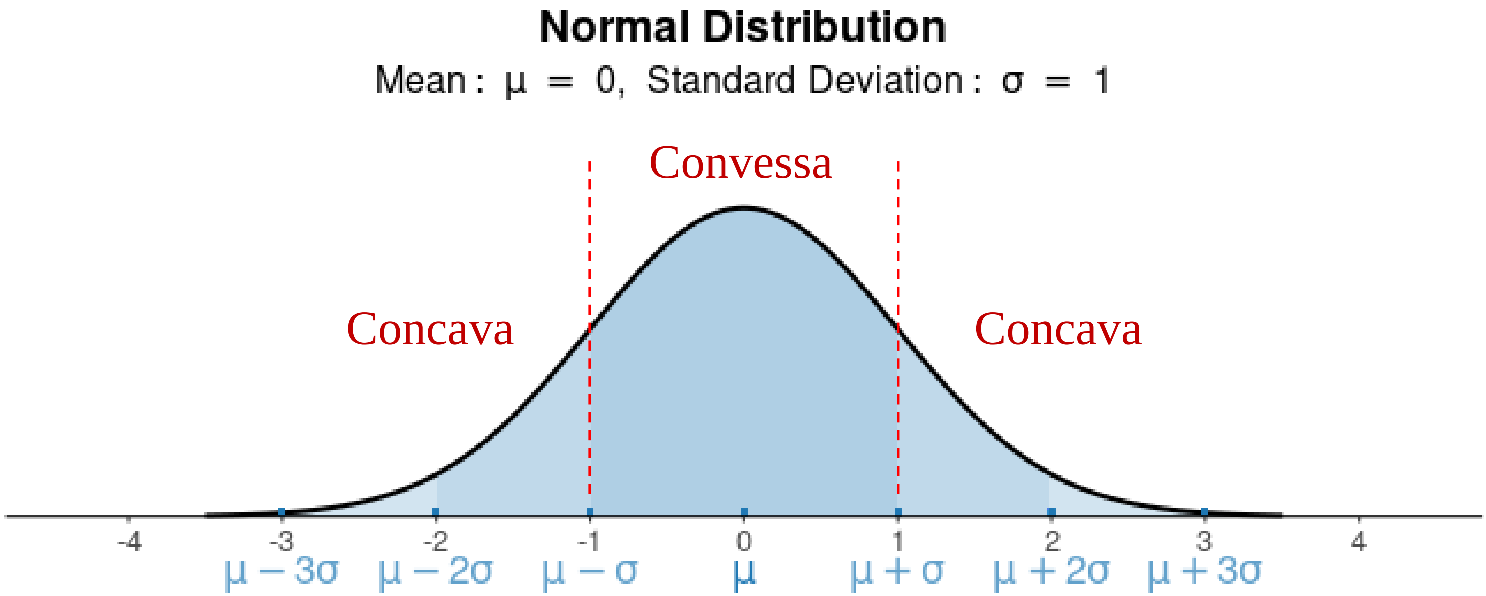
Spazio parametrico	$\Theta = \{(\mu, \sigma^2): \mu \in \mathbb{R};\ \sigma^2 \in \mathbb{R}_0^+\} \subset \mathbb{R}^2$

Semipiano positivo (chiuso) delle ordinate

Simboli μ e σ^2 impiegati usualmente per denotare media e varianza di una v.c. qualsiasi

Proprietà soddisfatte dalla funzione di densità	
$\frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right\} > 0$	$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right\} dx = 1$

$X \sim N(\mu, \sigma^2)$	$x \in \mathbb{R}$	$Z = \frac{X-\mu}{\sigma}$	$z = \frac{x-\mu}{\sigma} \in \mathbb{R}$
Come ogni v.c. **standardizzata**		$E(Z) = 0$	$V(Z) = 1 = SD(Z)$
La v.c. Z ha distribuzione di probabilità ancora **Normale** (di parametri 0 e 1) (Z v.c. Normale **standardizzata**)			
Funzione/Modello di **densità** parametrica/o della v.c. Z			$f(z; 0,1) = f(z)$
Notazione	$\phi(z) = \frac{1}{\sqrt{2\pi}} exp\left\{-\frac{1}{2}z^2\right\}$	$z \in \mathbb{R}$	$\phi(z) = \sigma \times f(x; \mu, \sigma^2)$ Funzioni di densità **proporzionali** per ogni valore $x \in \mathbb{R}$
$f(z) = \phi(z)$		$Z \sim N(0,1)$	
La v.c. Z ha distribuzione di probabilità **Normale standardizzata**			



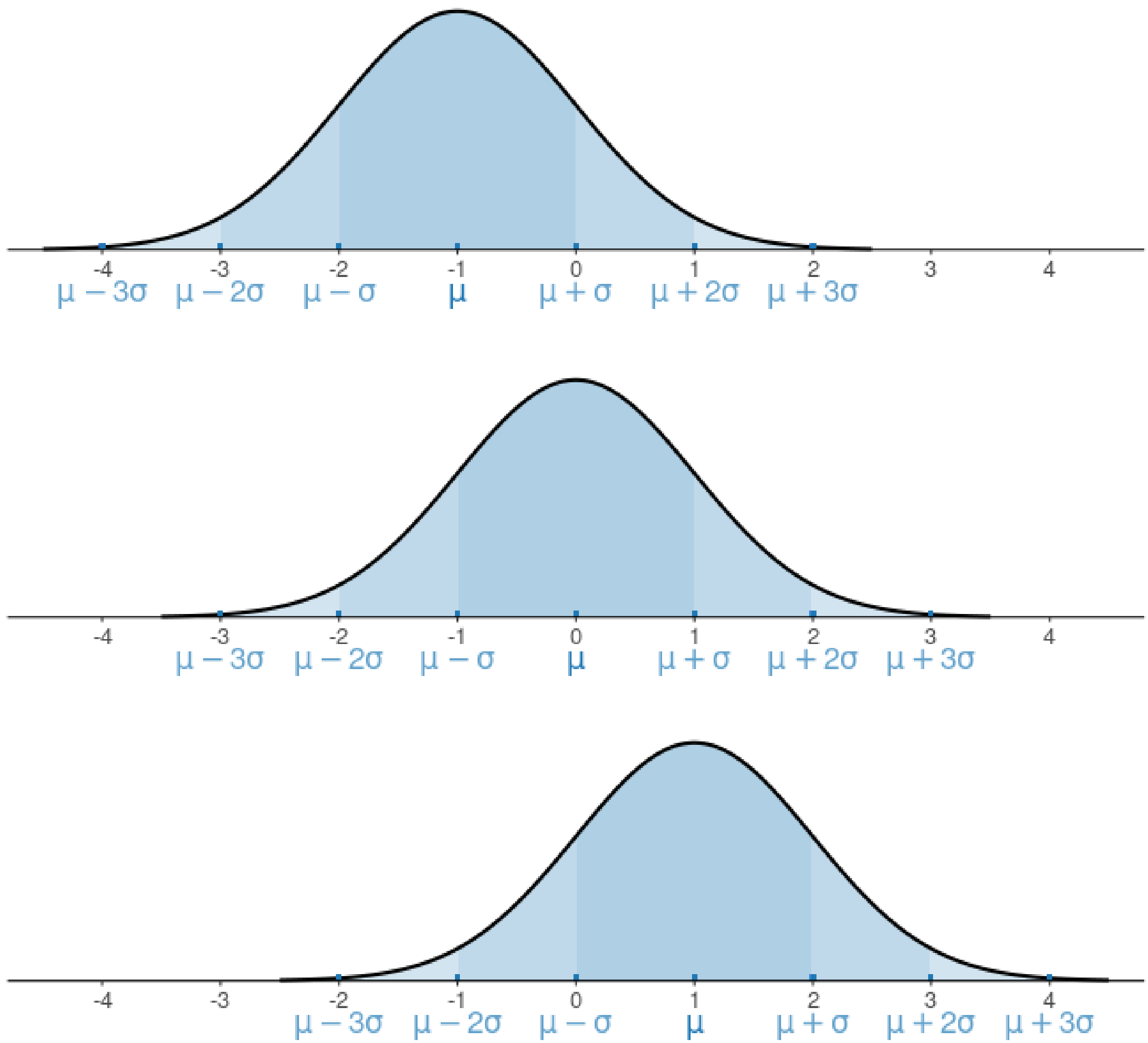
Studio analitico della funzione	Forma **campanulare simmetrica unimodale**

Simmetria rispetto al punto di ascissa $x = \mu$ $\quad \mu = E(X) = Me = Md \quad E(Z^3) = 0$

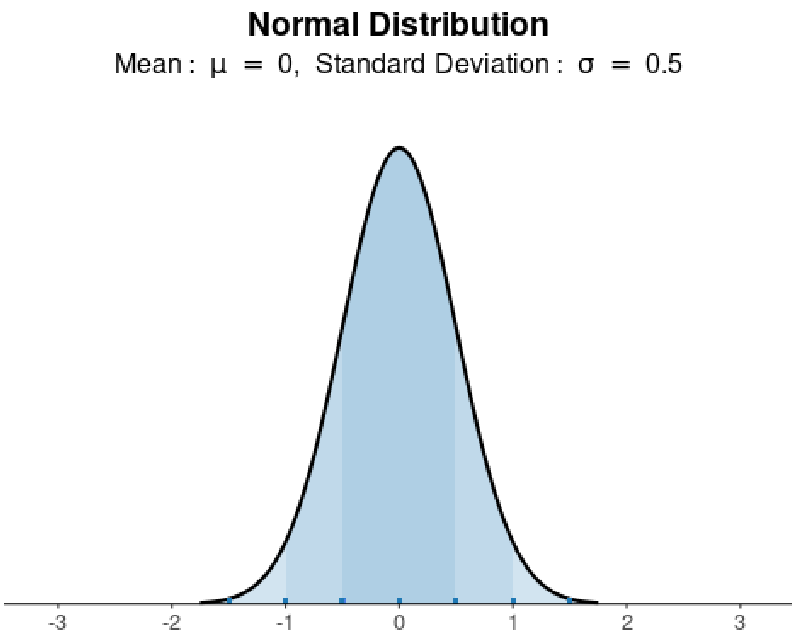
Asse x delle ascisse **asintoto orizzontale** $\quad \lim_{x \to \pm\infty} f(x; \mu, \sigma^2) = 0$

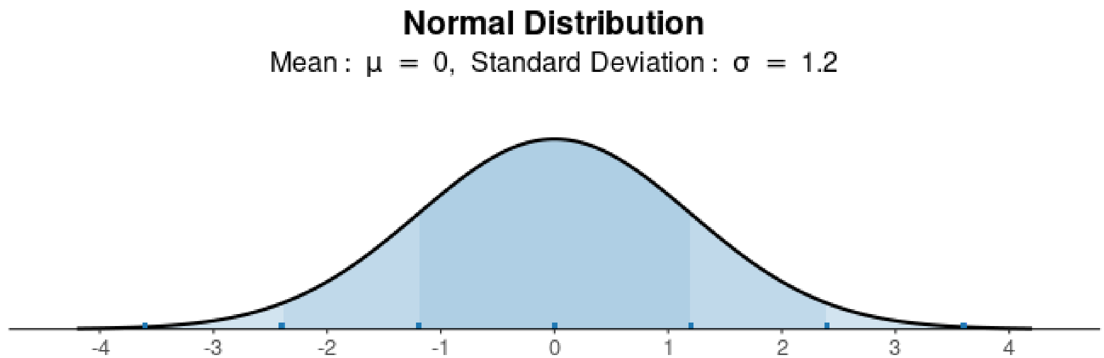
Due flessi *equidistanti* dal punto $x = \mu$ in corrispondenza delle ascisse $x = \mu \pm \sigma$

A parità di σ	La modifica di μ genera una **traslazione** della funzione $f(x;\mu,\sigma^2)$ lungo l'asse x

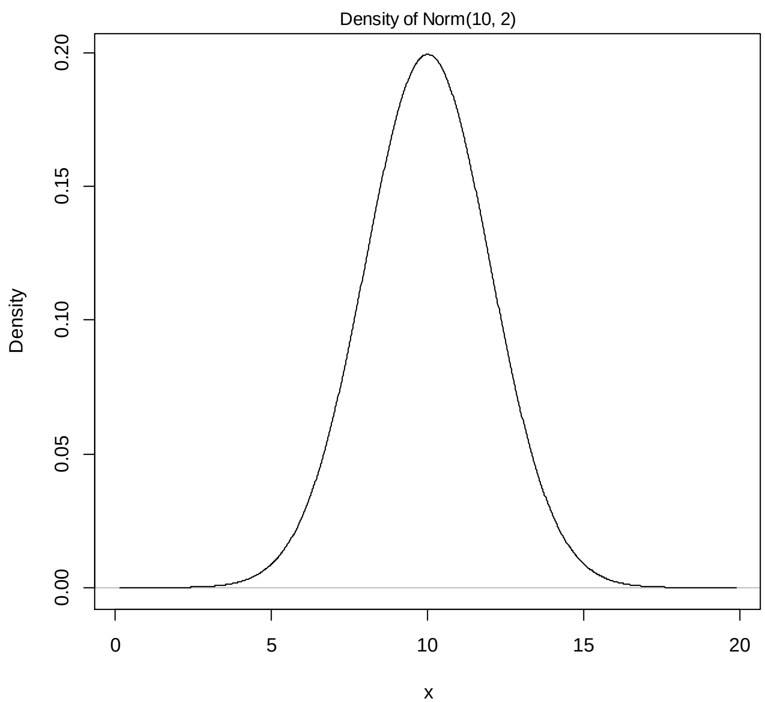


A parità di μ	All'**aumentare** di σ^2 i flessi si allontanano da μ	Maggiore probabilità a intervalli di valori **più distanti** da μ
	Al **diminuire** di σ^2 i flessi si avvicinano a μ	Maggiore probabilità a intervalli di valori **centrati** su μ
$\sigma^2 \to 0$	La distribuzione **tende** ad essere **degenere** con valori sempre più vicini a $x = \mu$	





Normal Distribution: Mean=10, Standard deviation=2




```
> characRV(Norm(mean=10, sd=2), charact=c("expectation", "s
                    expectation sd skewness kurtosis
Norm(mean=10,sd=2)           10  2        0        0
```

$X \sim N(\mu, \sigma^2)$	$E(Z^3) = 0$	$E(Z^4) - 3 = 0$
Coefficiente di **curtosi** (oltre che di asimmetria) **pari a 0**		

La **curtosi** indica quanto è più **appuntita** (o più **piatta**) la forma della distribuzione di una data v.c. rispetto a quella della v.c. Normale (che funge da v.c. di **riferimento**)

Di conseguenza, indica il **peso** più o meno accentuato delle **code** rispetto alla **parte centrale** della distribuzione, con riferimento alla forma di una v.c. Normale

Studio della curtosi utile per v.c. **unimodali** e di forma **simmetrica**

Ad esempio per la v.c. t di Student (ndr)

$X \sim N(\mu, \sigma^2)$	$x \in \mathbb{R}$	$Z = \dfrac{X-\mu}{\sigma} \sim N(0,1)$	$z = \dfrac{x-\mu}{\sigma} \in \mathbb{R}$

Per ogni $x \in \mathbb{R}$, e il corrispondente valore standardizzato $z \in \mathbb{R}$

Evento $(X \leq x)$ **equivalente** all'evento $(Z \leq z)$

$$F(x;\mu,\sigma^2) = P(X \leq x) = P(Z \leq z) = F(z; 0,1\ ) = F(z)$$

Notazione

$$F(z) = \Phi(z)$$

$F(x;\mu,\sigma^2)$ **funzione di ripartizione** di una v.c. Normale $X \sim N(\mu, \sigma^2)$

$\Phi(z)$ **funzione di ripartizione** della v.c. Normale **standardizzata** $Z \sim N(0,1)$

$$F(x;\mu,\sigma^2) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{1}{2}\frac{(t-\mu)^2}{\sigma^2}\right\} dt = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} exp\left\{-\frac{1}{2}t^2\right\} dt = \Phi(z)$$

La **funzione di ripartizione** di una **qualsiasi** v.c. Normale di parametri μ e σ^2 **coincide** con la **funzione di ripartizione** (unica) della corrispondente v.c. Normale **standardizzata**

$X \sim N(\mu, \sigma^2)$	Per ogni coppia $x_1, x_2 \in \mathbb{R}$	Intervallo reale $[x_1, x_2]$
	$x_1 < x_2$	Evento generico di interesse

$$P(x_1 \leq X \leq x_2) = F(x_2; \mu, \sigma^2) - F(x_1; \mu, \sigma^2) = \Phi(z_2) - \Phi(z_1) = P(z_1 \leq Z \leq z_2)$$

$z_1 = \dfrac{x_1 - \mu}{\sigma}$	$z_2 = \dfrac{x_2 - \mu}{\sigma}$	Estremi standardizzati dell'intervallo $[x_1, x_2]$	$[z_1, z_2]$ intervallo reale dei valori standardizzati

Per calcolare la probabilità che una **qualsiasi** v.c. Normale di parametri μ e σ^2 assuma valori in un **qualsiasi** intervallo reale **è sufficiente standardizzare** gli estremi dell'intervallo e ricondursi all'uso della **funzione di ripartizione** della v.c. Normale **standardizzata**

$$P(x_1 \leq X \leq x_2) = \Phi\left(\frac{x_2 - \mu}{\sigma}\right) - \Phi\left(\frac{x_1 - \mu}{\sigma}\right)$$

La funzione $\Phi(\cdot)$ **non** ha una forma analitica *esplicita* (ovviamente, neanche $F(\cdot\,; \mu, \sigma^2)$)

Valori (accurati) di $\Phi(z)$ calcolati mediante algoritmi di *integrazione numerica*

Forniti nei software di analisi statistica e nei linguaggi di programmazione

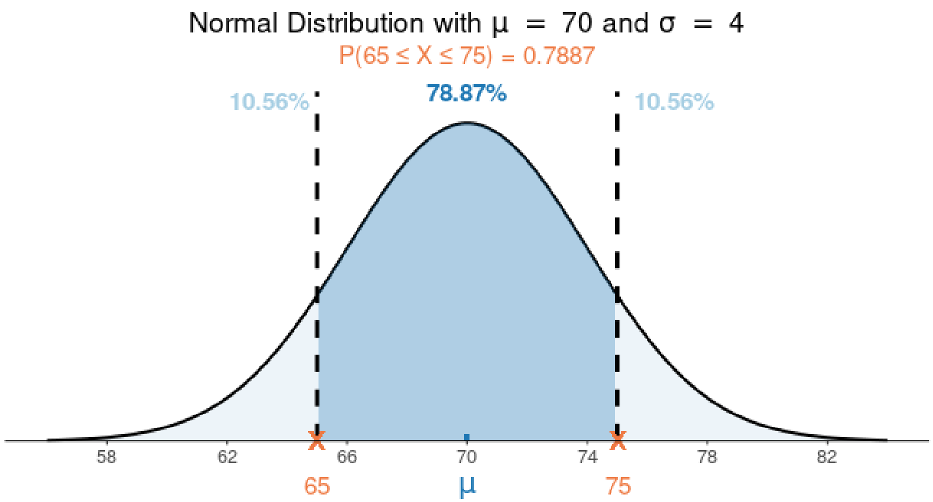
Anche presenti in **tavole** appositamente predisposte per i valori z più rilevanti

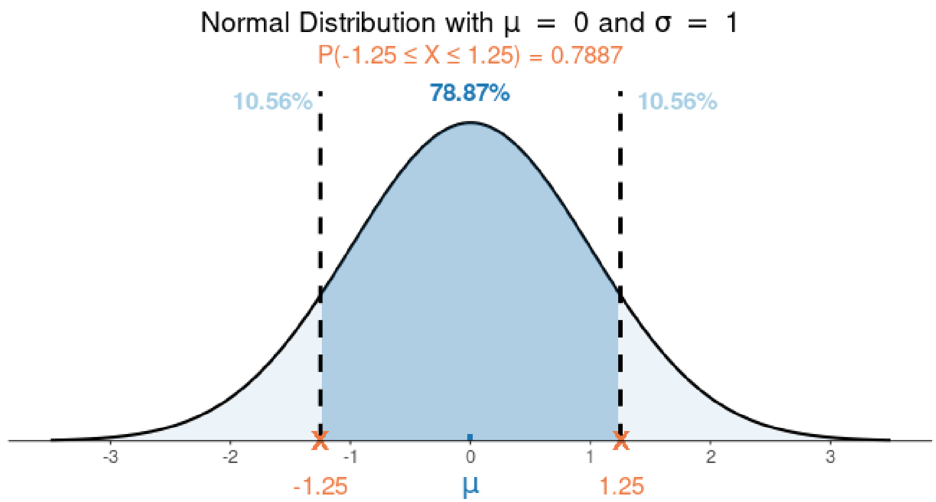
$X \sim N(70,16)$	Ad esempio, X è il **peso** di persone adulte	Ipotesi di Normalità appropriata?
Peso v.c. **non negativa** e con supporto **limitato**!		

Impossibile osservare valori di peso tanto **distanti** dalla **media** (il centro della distribuzione)	Tali valori generano eventi (intervalli di peso) **impossibili** (di probabilità 0)

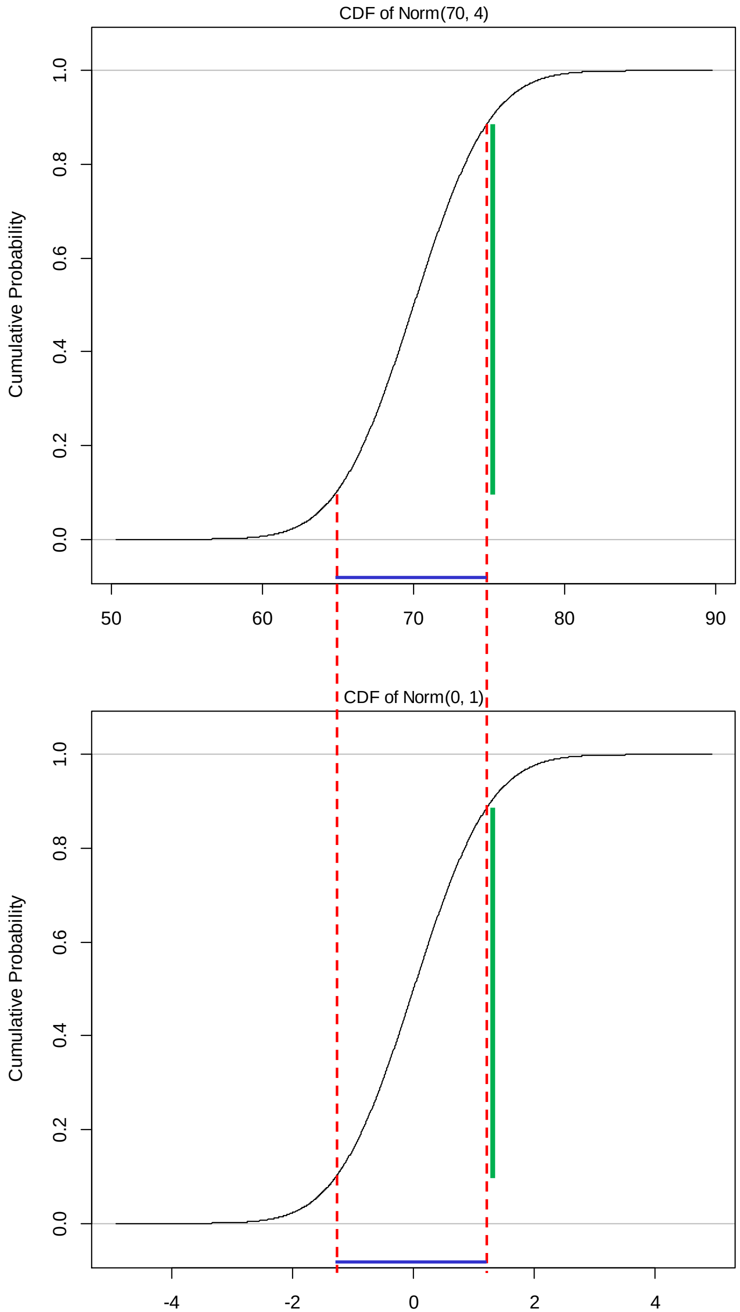
Si può pensare che abbiano probabilità infinitesimale, pari ad esempio a 10^{-20} (di fatto pari a 0)

Distribuzione, di fatto, totalmente **addensata** nella sua **parte centrale**	**Normalità ragionevole**
Oltre che di forma **campanulare simmetrica**	





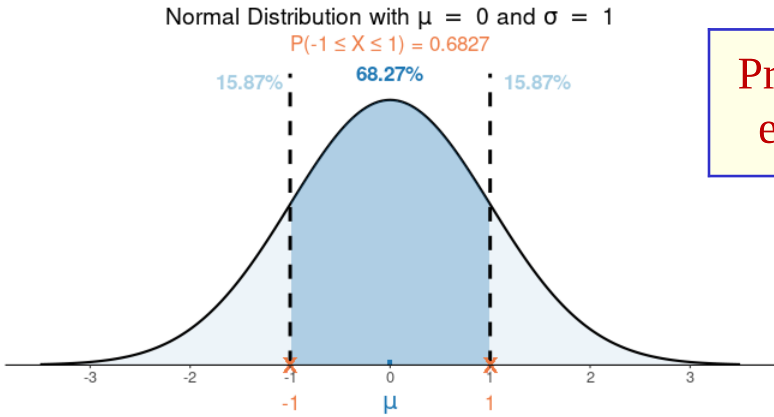
Curve di **densità** perfettamente **sovrapponibili**

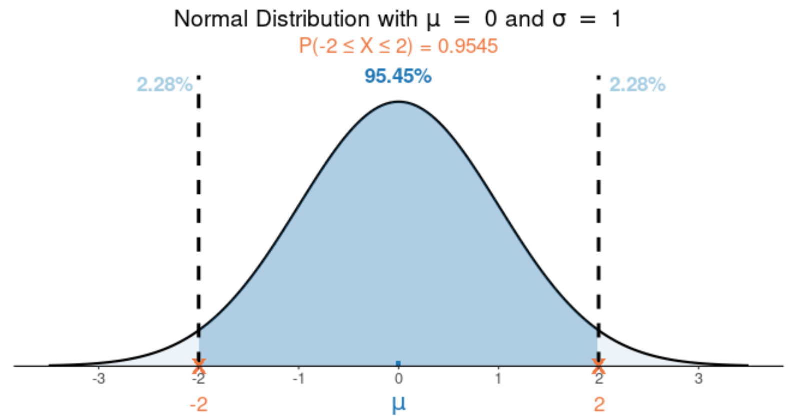


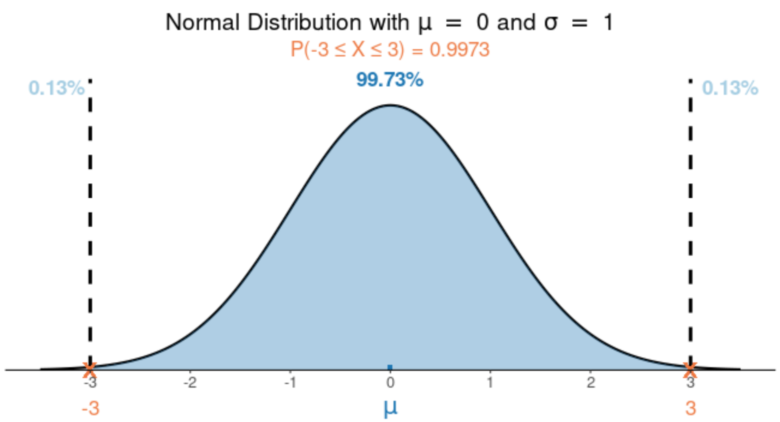
Curve di **ripartizione** (ogive) perfettamente **sovrapponibili**

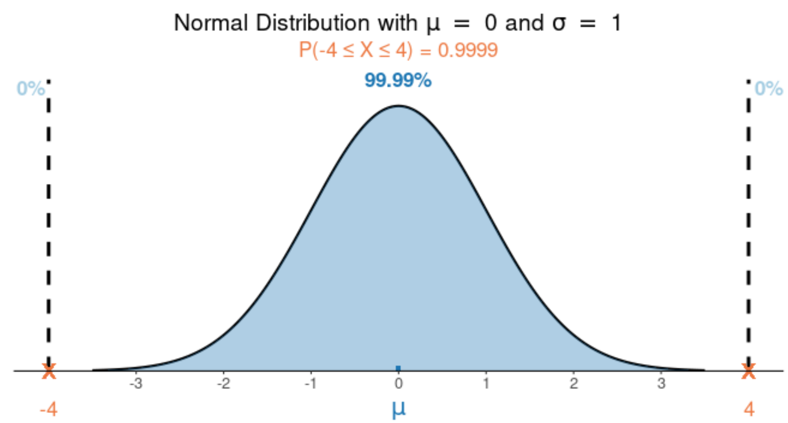
Probabilità di intervalli centrati su μ e da essa distanti per multipli di σ

Per una qualsiasi v.c. $X \sim N(\mu, \sigma^2)$



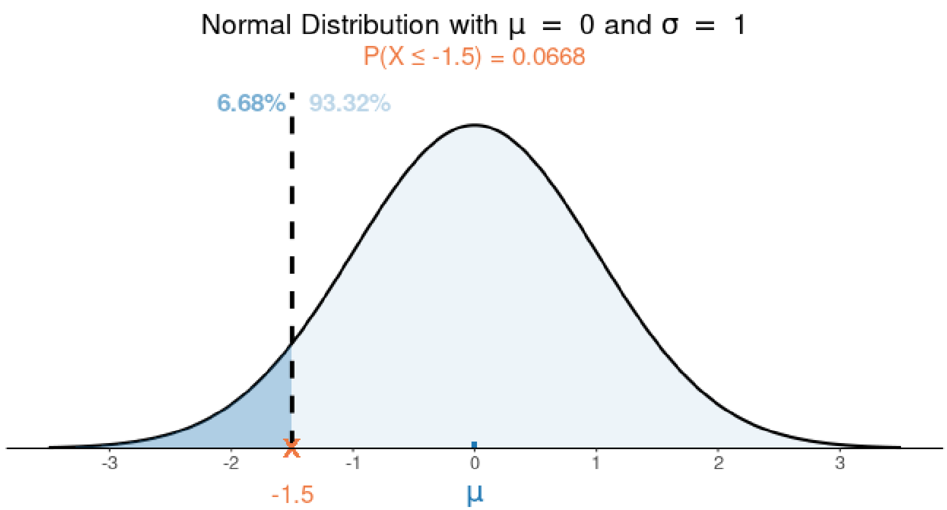


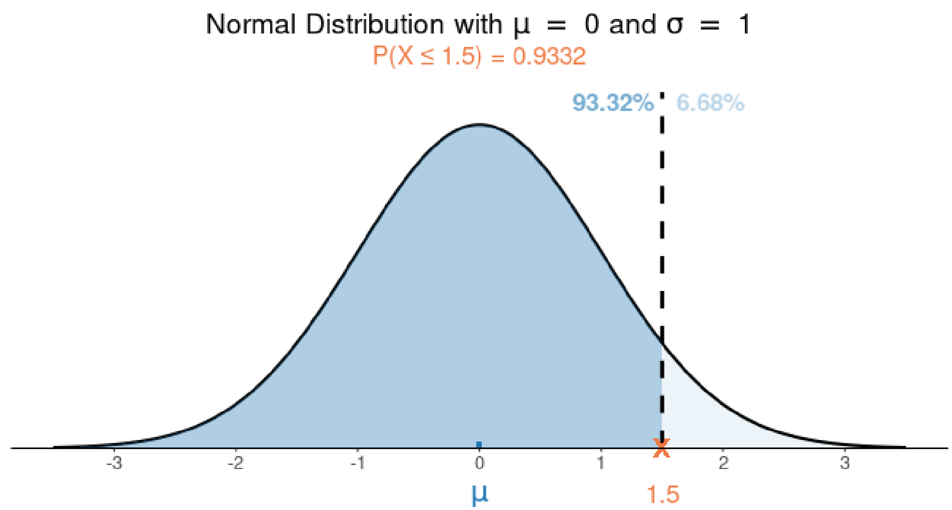




Valori di v.c. Normali più **lontani** da μ del **doppio** o del **triplo** di σ

Valori anomali





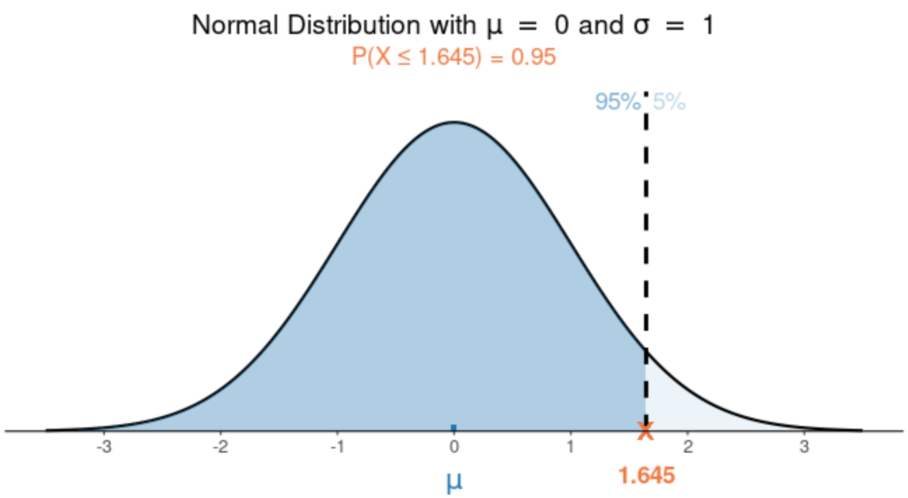
Per ogni valore **positivo** $z > 0$	$\Phi(-z) = 1 - \Phi(z)$

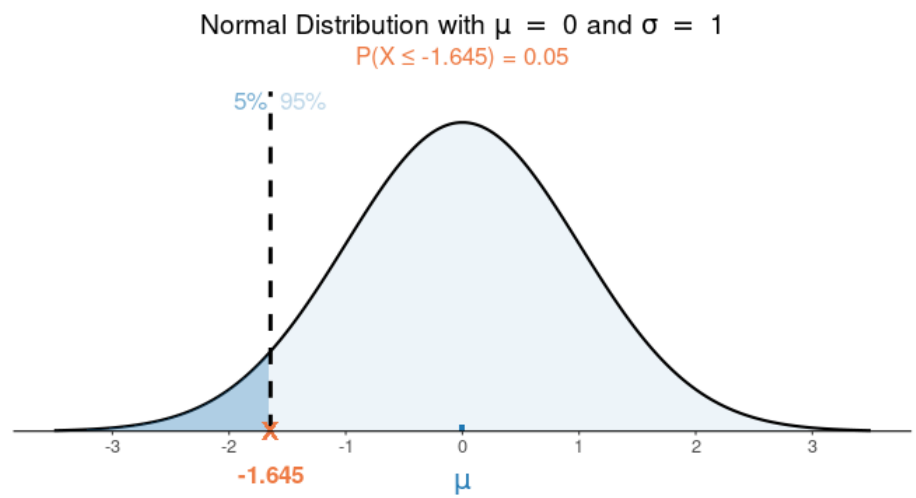
A probabilità **complementari** corrispondono quantili **opposti**

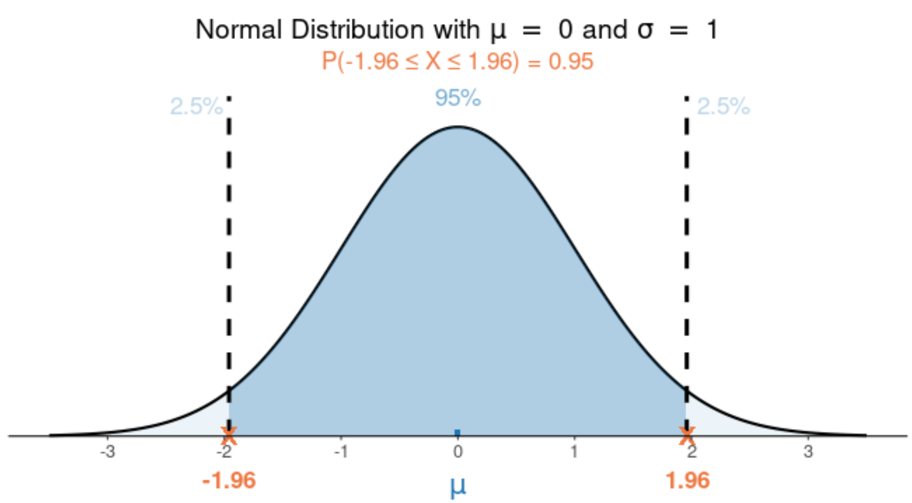
$\Phi(z)$ **tabulata** solo per valori di z positivi

Ovviamente, $\Phi(0) = 0.5$

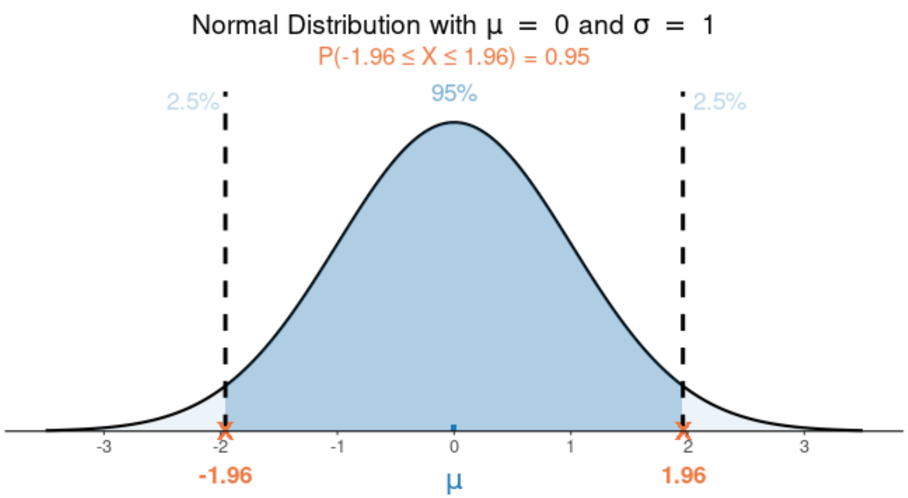
Quantile di livello p di una v.c. $Z \sim N(0,1)$	Ovviamente non calcolabile in modo esplicito	
Valore z_p tale che	$\Phi(z_p) = p$	$1 - \Phi(z_p) = 1 - p$



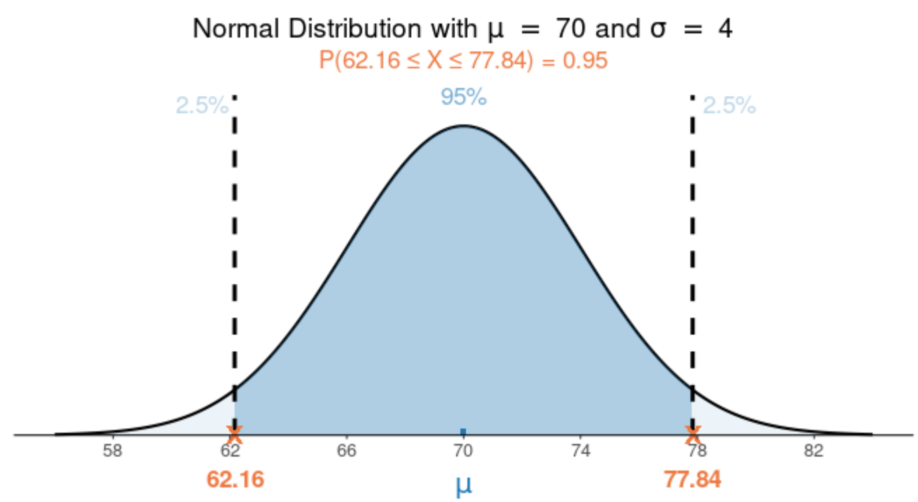




Quantile x_p di livello p di una qualsiasi v.c. $X \sim N(\mu, \sigma^2)$		
$z_p = \dfrac{x_p - \mu}{\sigma}$	z_p valore standardizzato del quantile x_p	$x_p = \mu + z_p \sigma$



$$X \sim N(70, 16)$$



$$x_{0.025} = 70 - 1.96 \times 4 = 62.16$$

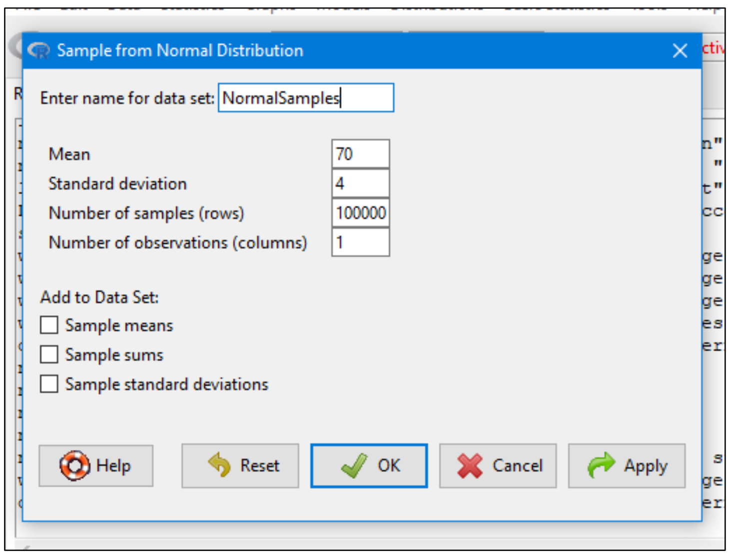
$$x_{0.975} = 70 + 1.96 \times 4 = 77.84$$

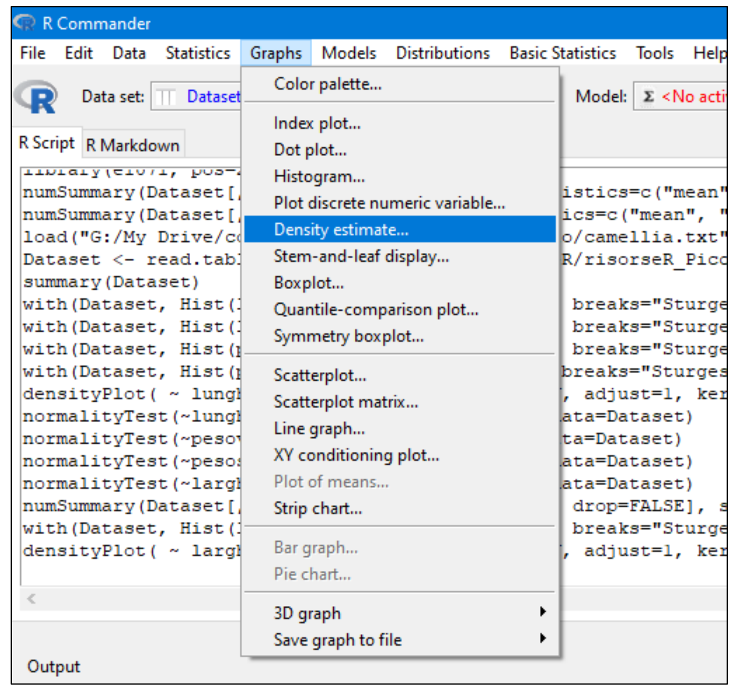
Logica frequentista

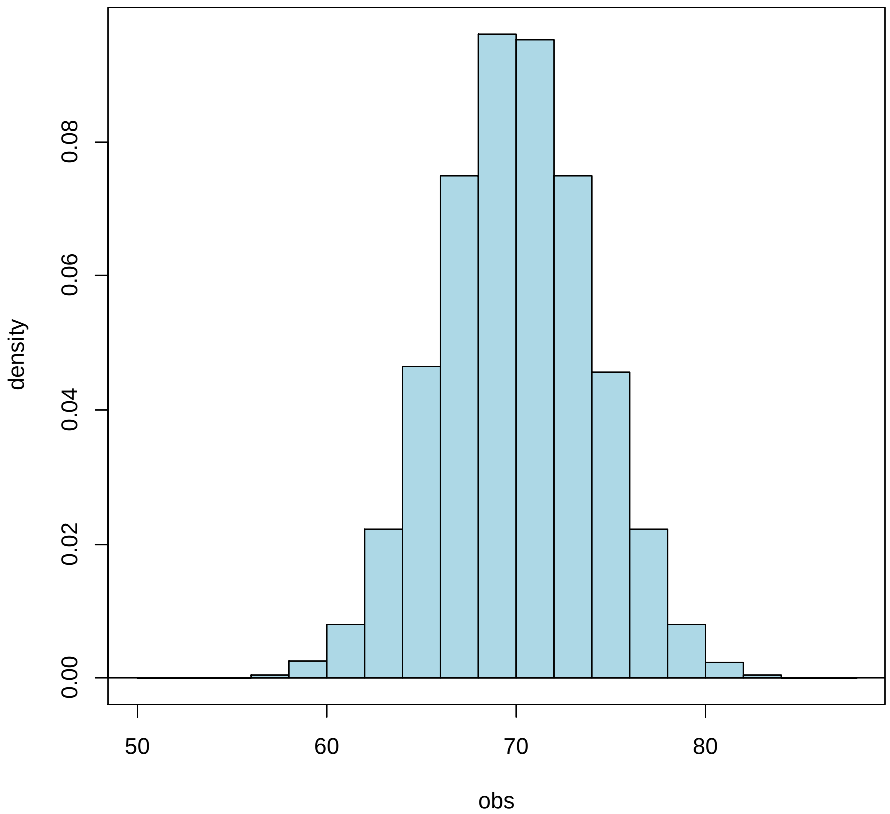
Distribuzione empirica di una v.c. $X \sim N(70,16)$

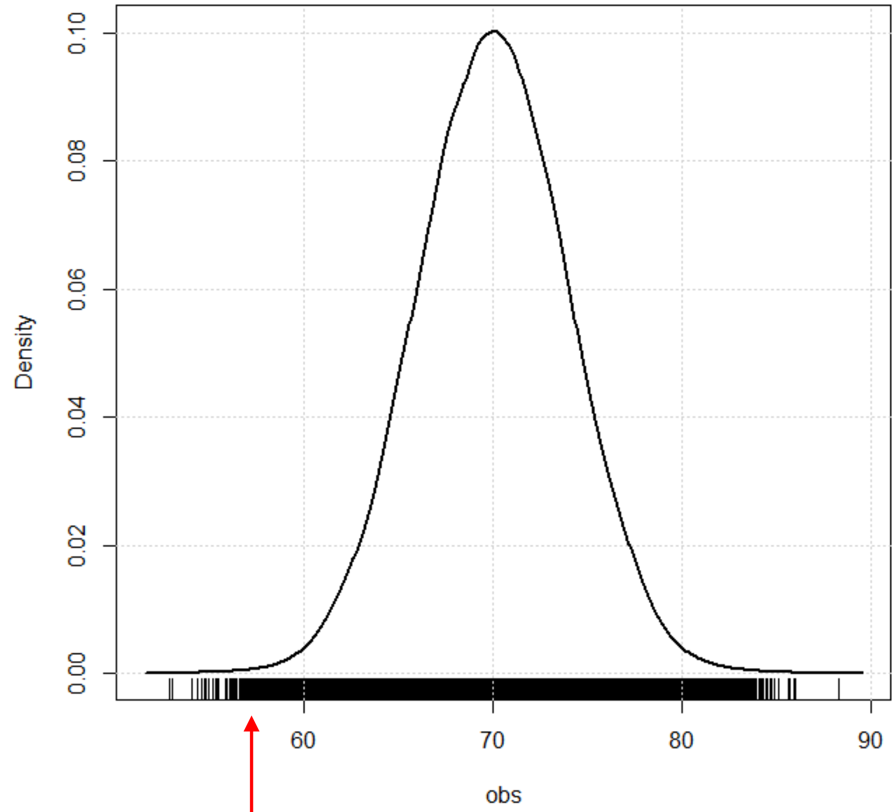
Istogramma e (meglio) **density plot**

Sintesi numeriche







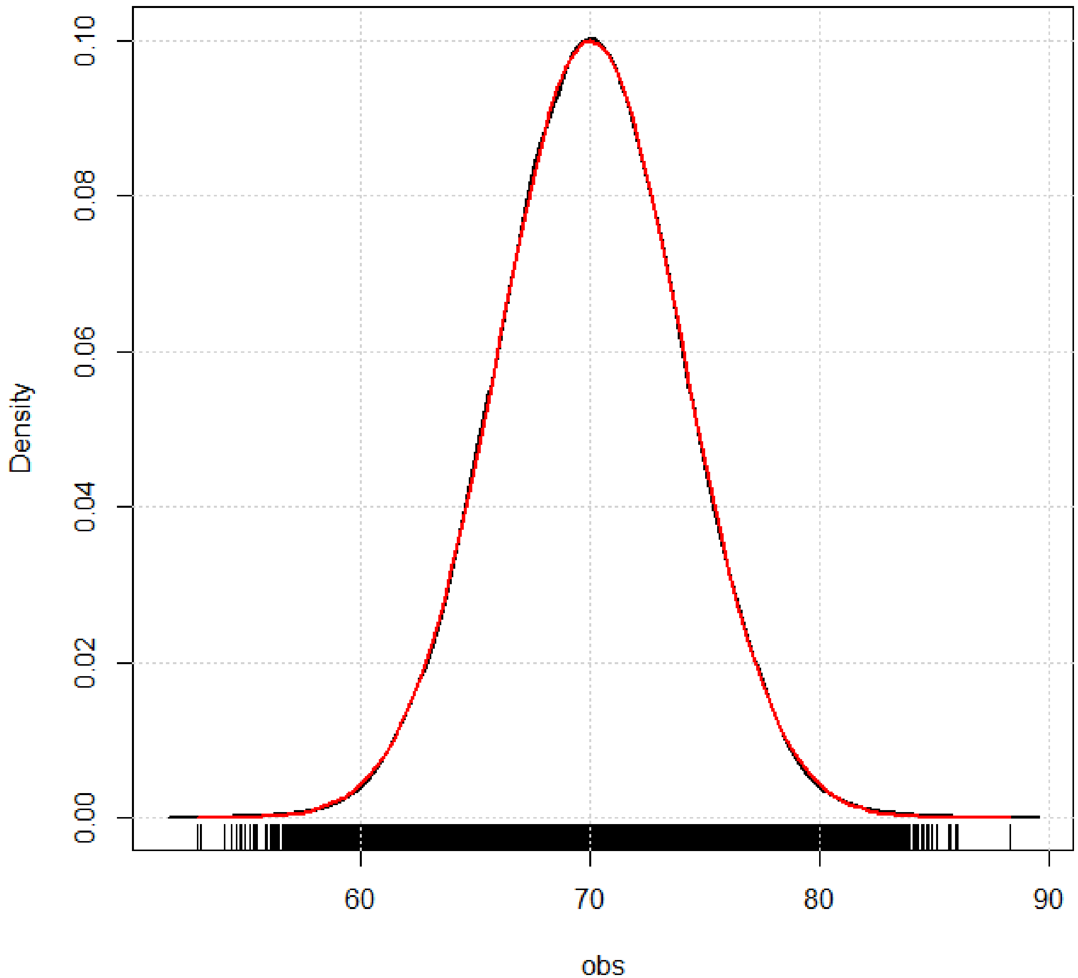



```
Numerical summary:
    mean       sd      skewness     kurtosis      50%      n
 69.98592 3.994239 -0.001099112 -0.008033291 69.98305 100000
```

Rug plot Singoli valori "marcati" sull'asse delle ascisse

```
#
# Distribuzione empirica vs distribuzione di probabilità - Modelli continui
#
# Generazione in Rcmdr di m campioni di ampiezza 1 da una popolazione continua
# Ad esempio, 100000 campioni di ampiezza 1 da una popolazione N(70,16)
#
NormalSamples <- as.data.frame(matrix(rNorm(mean=70, sd=4))(100000*1), ncol=1))
rownames(NormalSamples) <- paste("sample", 1:100000, sep="")
colnames(NormalSamples) <- "obs"
#
# Density plot in Rcmdr
#
densityPlot( ~ obs, data=NormalSamples, bw=bw.SJ, adjust=1, kernel=dnorm, method="adaptive")
#
#
data <- NormalSamples      # inserire nome del dataframe    #
attach(data)                                                #
x <- seq( min(obs), max(obs), length=100)                   #
#                                                           #
lines( x, dnorm(x, mean(obs),sd(obs)), lwd=2, col="red")    # lanciare questo comando per tracciare la curva teorica - Normale
#                                                           #
detach(data)                                                #
#
```

Confronto tra funzione di densità empirica e funzione di densità teorica



Proprietà riproduttiva della v.c. Normale

$X_1, X_2, \cdots, X_n$ v.c. Normali e **indipendenti**				Non necessariamente identicamente distribuite
$X_i \sim N(\mu_i, \sigma_i^2)$	$\mu_i = E(X_i)$	$\sigma_i^2 = V(X_i)$	$i = 1,2,\cdots,n$	
$Y = a_1X_1 + a_2X_2 + \cdots + a_nX_n$			Combinazione lineare delle n v.c.	
Una **combinazione lineare** di v.c. Normali e **indipendenti** (la v.c. Y) è ancora una v.c. **Normale**				
$Y \sim N(\mu_0, \sigma_0^2)$	$\mu_0 = \sum_i a_i\mu_i$	$\sigma_0^2 = \sum_i a_i^2\sigma_i^2$	Cfr Parte 4, pagg 11-12	
$X_1, X_2, \cdots, X_n$ v.c. Normali **indipendenti** e **identicamente distribuite** (v.c. IID)				
$X_i \sim N(\mu, \sigma^2)$	$i = 1,2,\cdots,n$	$Y \sim N(\mu_0, \sigma_0^2)$	$\mu_0 = \mu\sum_i a_i$	$\sigma_0^2 = \sigma^2\sum_i a_i^2$
Casi speciali (v.c. IID)				
Somma di v.c. Normali IID		$S = \sum X_i \sim N(n\mu, n\sigma^2)$		
Media aritmetica di v.c. Normali IID		$\overline{X} = S/n \sim N(\mu, \sigma^2/n)$		

Teorema del limite centrale	Basato su una **evidenza empirica** antica e costante

Prime formalizzazioni	Formalizzazione più nota
De Moivre (1718 e 1733), Gauss (1809), Laplace (dal 1812)	Lindeberg-Lévy (1922 e 1925)
	Estesa in varie direzioni

$(X_1, X_2, \cdots)$ **successione** (insieme infinito) di v.c. **indipendenti** e **identicamente distribuite**

$X_1, X_2, \cdots$ v.c. discrete o continue di tipo **qualsiasi** (anche ignoto) con **media** e **varianza** entrambe *finite*	$X_n, n = 1,2,\cdots$	
	$E(X_n) = \mu$	$V(X_n) = \sigma^2$

$(S_1, S_2, \cdots)$ **successione** corrispondente delle v.c. **somme**

$S_n = X_1 + X_2 + \cdots + X_n = \sum X_i$	$S_1 = X_1$	$S_3 = X_1 + X_2 + X_3$	$E(S_n) = n\mu$
	$S_2 = X_1 + X_2$	$\vdots$	$V(S_n) = n\sigma^2$

$(Z_1, Z_2, \cdots)$ **successione** delle v.c. **somme standardizzate**

$Z_n = \dfrac{S_n - n\mu}{\sqrt{n\sigma^2}}$	$Z_1 = \dfrac{S_1 - 1\mu}{\sqrt{1\sigma^2}}$	$Z_2 = \dfrac{S_2 - 2\mu}{\sqrt{2\sigma^2}}$	$Z_3 = \dfrac{S_3 - 3\mu}{\sqrt{3\sigma^2}}$	$\cdots$	$E(Z_n) = 0$
					$V(Z_n) = 1$

$$Z_n = \frac{S_n - n\mu}{\sqrt{n\sigma^2}} = \frac{\overline{X}_n - \mu}{\sqrt{\sigma^2/n}}$$

$E(Z_n) = 0$

$V(Z_n) = 1$

$$\overline{X}_n = \frac{S_n}{n} = \frac{1}{n}\sum X_i$$

$E(\overline{X}_n) = \mu$

$V(\overline{X}_n) = \sigma^2/n$

La v.c. **somma standardizzata coincide** con la v.c. **media aritmetica standardizzata**

$(Z_1, Z_2, \cdots)$ **successione** delle v.c. **somme standardizzate** o **medie standardizzate**

Z_n ha distribuzione di probabilità (nota o ignota) *derivata* da quella di X_n (nota o ignota)

Il **teorema del limite centrale** (TLC) fornisce
la distribuzione di probabilità di Z_n quando $n \to +\infty$

Distribuzione **limite** di Z_n

Le v.c. **somma standardizzata** e **media standardizzata** di v.c. IID (con **media** e **varianza** finite) hanno distribuzione **limite Normale standardizzata**

Convergono in **distribuzione** alla v.c. **Normale standardizzata**

$$Z_n = \frac{S_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{d} N(0{,}1)$$

$$Z_n = \frac{\overline{X}_n - \mu}{\sqrt{\sigma^2/n}} \xrightarrow{d} N(0{,}1)$$

Risultato valido per v.c. $X_1, X_2, \cdots$ di tipo **qualsiasi** (anche ignoto)

Teorema del limite centrale		**Implicazioni**
Più è **grande** il valore di n, più la distribuzione di probabilità della v.c. Z_n si **avvicina** alla distribuzione **Normale standardizzata**		
Per n **sufficientemente grande**		
La v.c. Z_n (**somma standardizzata** o **media standardizzata)** ha distribuzione **asintotica** (cioè approssimata) **Normale standardizzata**		
$Z_n = \dfrac{S_n - n\mu}{\sqrt{n\sigma^2}} \sim N(0,1)$	$Z_n = \dfrac{\overline{X}_n - \mu}{\sqrt{\sigma^2/n}} \sim N(0,1)$	In modo **approssimato**
Le v.c. S_n (**somma**) e $\overline{X}_n$ (**media aritmetica**) hanno distribuzioni **asintotiche Normali** (con i rispettivi parametri)		
$S_n \sim N(n\mu, n\sigma^2)$	$\overline{X}_n \sim N(\mu, \sigma^2/n)$	In modo **approssimato**
Nella **logica inferenziale**, interesse **speciale** per la **media** $\overline{X}_n$ e la **media standardizzata** Z_n		

Teorema del limite centrale		**Implicazioni**
Più è **grande** il valore di n, più la distribuzione di probabilità della v.c. Z_n si **avvicina** alla distribuzione **Normale standardizzata**		
Per n **sufficientemente grande**		
La v.c. Z_n (**somma standardizzata** o **media standardizzata)** ha distribuzione **asintotica** (cioè approssimata) **Normale standardizzata**		
$Z_n = \dfrac{S_n - n\mu}{\sqrt{n\sigma^2}} \sim N(0,1)$	$Z_n = \dfrac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} \sim N(0,1)$	In modo **approssimato**
Le v.c. S_n (**somma**) e $\bar{X}_n$ (**media aritmetica**) hanno distribuzioni **asintotiche Normali** (con i rispettivi parametri)		
$S_n \sim N(n\mu, n\sigma^2)$	$\bar{X}_n \sim N(\mu, \sigma^2/n)$	In modo **approssimato**
Nella **logica inferenziale**, interesse **speciale** per la **media** $\bar{X}_n$ e la **media standardizzata** Z_n		

Caso speciale del teorema del limite centrale	$X_1, X_2, \cdots$ v.c. di **Bernoulli**	
	$E(X_n) = p$	$V(X_n) = p(1-p)$
$S_n = \sum X_i$	**Numero** di successi nell'esperimento binomiale	
	$E(S_n) = np$	$V(S_n) = np(1-p)$
$\bar{X}_n = \frac{1}{n}\sum X_i$	**Proporzione** di successi nell'esperimento binomiale	
	$E(\bar{X}_n) = p$	$V(\bar{X}_n) = p(1-p)/n$

Le v.c. S_n e $\bar{X}_n$ hanno distribuzioni **note** per ogni n (sono entrambe v.c. **Binomiali**)

$S_n \sim Bin(n,p)$	$\bar{X}_n \sim (1/n)Bin(n,p)$	Cfr Parte 5

Tuttavia, per il TLC (noto in questo caso come **teorema di De Moivre-Laplace**)

$Z_n = \dfrac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow{d} N(0,1)$	$Z_n = \dfrac{\bar{X}_n - p}{\sqrt{p(1-p)/n}} \xrightarrow{d} N(0,1)$

Per n **sufficientemente grande**		In modo **approssimato**
$S_n \sim N(np, np(1-p))$	$\bar{X}_n \sim N(p, p(1-p)/n)$	$Z_n \sim N(0,1)$

Nella **logica inferenziale**, distribuzione **asintotica Normale**
di $\bar{X}_n$ e Z_n preferita a quella **esatta Binomiale**

Condizioni (in ordine di priorità) che **accelerano** la **convergenza** alla **Normalità**

Distribuzione di X_n (v.c. della successione genitrice)	▪ Unimodale ▪ Simmetrica

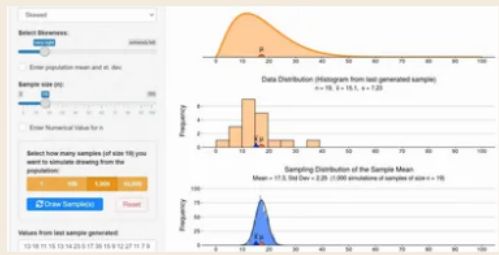
Naturalmente, una distribuzione *plurimodale* e *asimmetrica* ritarda la convergenza

Sampling Distributions and the Central Limit Theorem

Sampling Distribution of the Sample Proportion

Experience how the sampling distribution of the **sample proportion** builds up one sample at a time. Use sliders to explore the shape of the sampling distribution as the sample size n increases, or as the population proportion p changes. Overlay a normal distribution to explore the Central Limit Theorem.

Sampling Distribution of the Sample Mean (Continuous Population)



Experience how the sampling distribution of the **sample mean** builds up one sample at a time. Use a variety of real or theoretical **continuous** population distributions (or create your own) to draw samples from. Move sliders to explore when the Central Limit Theorem kicks in.

Sampling Distribution of the Sample Mean (Discrete Population)

Experience how the sampling distribution of the **sample mean** builds up one sample at a time. Use a variety of real or theoretical **discrete** population distributions (or create your own) to draw samples from. Move sliders to explore when the Central Limit Theorem kicks in.

Oltre alle implicazioni sulla distribuzione asintotica di **speciali** v.c. (somma e media aritmetica, anche standardizzate) il TLC ha un'altra **importante implicazione**

Tutte le volte che una **variabile casuale** continua (interprete di un fenomeno reale) può essere ricondotta alla **somma** (o media, o anche combinazione lineare) di un **gran numero** di **cause** (v.c.) **indipendenti** (nessuna delle quali *prevale* sulle altre)

Distribuzione di probabilità di tale variabile casuale ragionevolmente **approssimabile** dalla distribuzione **Normale** (che è dunque il modello di probabilità)

Per questa ragione, molte v.c. (ma non tutte!) sono **ben definite** come v.c. **Normali**

Nella **logica inferenziale**, sono v.c. di popolazione	**Esempio**	Peso di adulti

Determinato dall'effetto **additivo** (combinazione) di una **molteplicità** di fattori (cause) **indipendenti**, quali tra gli altri:

- il peso dei genitori e dei nonni (paterni e materni)
- l'alimentazione durante la gestazione e la prima infanzia
- le abitudini alimentari attuali
- eventuali traumi o malattie
- l'abitudine all'attività sportive
- la latitudine in cui si è nati

Variabile casuale t **di Student**

William Sealy **Gosset** (1876 – 1937)

Pseudonimo "Student"

Variabile casuale **campionaria** introdotta in un articolo del 1908

È una speciale **derivazione** di una v.c. **Normale standardizzata**

Usualmente denominata con T

Caratterizzata dal **parametro** g	g numero intero ≥ 1

Il parametro g è denominato **gradi di libertà** (gl)

$T \sim t(g)$	La v.c. T ha distribuzione di probabilità t **di Student** con g gl

$T \sim t(g)$	**Variabile casuale** continua che *può assumere* come valore un **qualsiasi** numero reale
$t \in \mathbb{R} = (-\infty, \infty)$ $\mathbb{R}$ **supporto** della v.c.	La v.c. T (t di Student) "eredita" il supporto (ma non solo!) della v.c. Z (Normale standardizzata)

L'**insieme** delle v.c. *generato* da *tutti i possibili* valori di g (sono una *infinità numerabile*) è la **famiglia parametrica** di v.c.

Funzione/Modello di **densità** parametrica/o di una v.c. $T \sim t(g)$

$$f(t; g) = \frac{1}{\sqrt{g\pi}} \frac{\Gamma\left(\frac{g+1}{2}\right)}{\Gamma\left(\frac{g}{2}\right)} \left(1 + \frac{t^2}{g}\right)^{-\frac{g+1}{2}}$$

$t \in (-\infty, \infty)$	**Supporto**	$g \in \{1, 2, \cdots, \infty\}$	**Spazio parametrico**

L'**insieme** delle $f(t; g)$ *generato* dai valori di g è la **famiglia parametrica** di funzioni di **densità** (modelli di probabilità) di tipo t **di Student** con g gradi di libertà

Media e **varianza** di una v.c. $T \sim t(g)$		
$g \geq 2$	$E(T) = 0$	La media **non esiste** per $g = 1$
$g \geq 3$	$1 < V(T) = \dfrac{g}{g-2} \leq 3$	La varianza **non esiste** per $g = 1$ e $g = 2$
$g = 1$		**Non esistono** nè media nè varianza
$g = 2$		**Esiste** la media, **non esiste** la varianza
$g \geq 3$		**Esistono** media e varianza
$g \geq 2$	La v.c. T (t di Student) "eredita" la media (di valore 0) della v.c. Z (Normale standardizzata)	
$g \geq 3$	$\lim\limits_{g \to \infty} \left(V(T) = \dfrac{g}{g-2} \right) = 1$	All'aumentare di g, $V(T)$ è sempre più prossima a $V(Z) = 1$

Per ogni valore di g	$Me = Md = 0$

Distribuzione di forma **campanulare simmetrica** (rispetto a 0) **unimodale**

$g \geq 2$	$E(T) = Me = Md = 0$	$g \geq 4$	$E(Z^3) = 0$

Asse t delle ascisse **asintoto orizzontale**	$\lim_{t \to \pm\infty} f(t; g) = 0$

Due flessi *equidistanti* dal punto $t = 0$ in corrispondenza di $t = \pm\sqrt{g/(g+2)}$

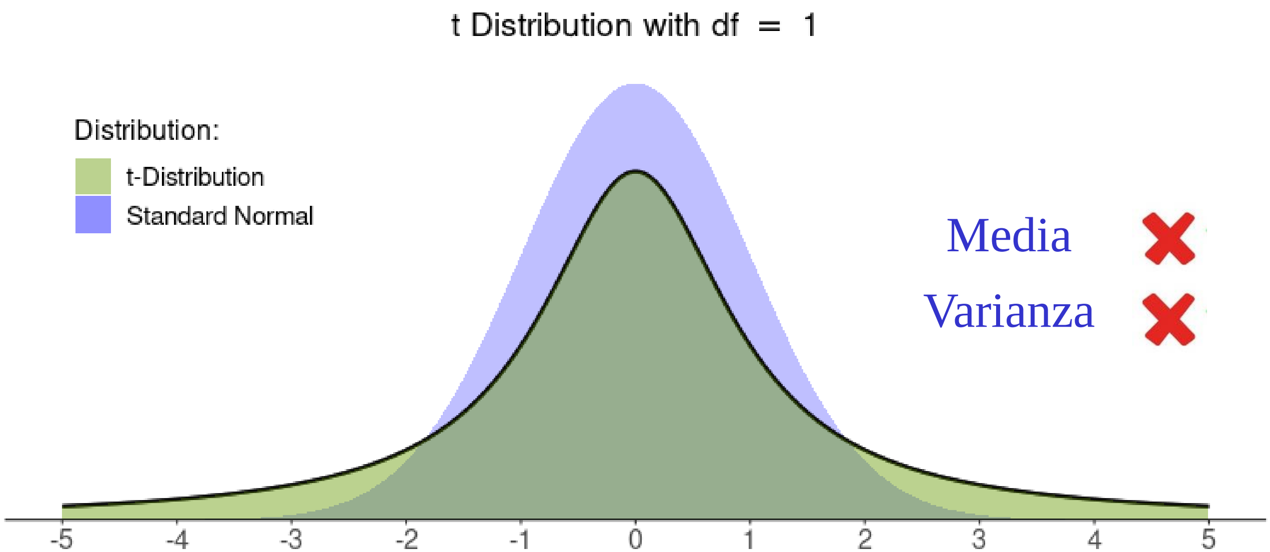
$-1 \leq -\sqrt{g/(g+2)} \leq 0$	$\xrightarrow[g \to \infty]{} -1$	$0 \leq +\sqrt{g/(g+2)} \leq +1$	$\xrightarrow[g \to \infty]{} +1$

Forma **simile** a quella di una distribuzione **Normale standardizzata**

Per $g \to \infty$	$T \xrightarrow{d} N(0,1)$	**In virtù di risultati "limite" più estesi rispetto al solo TLC**

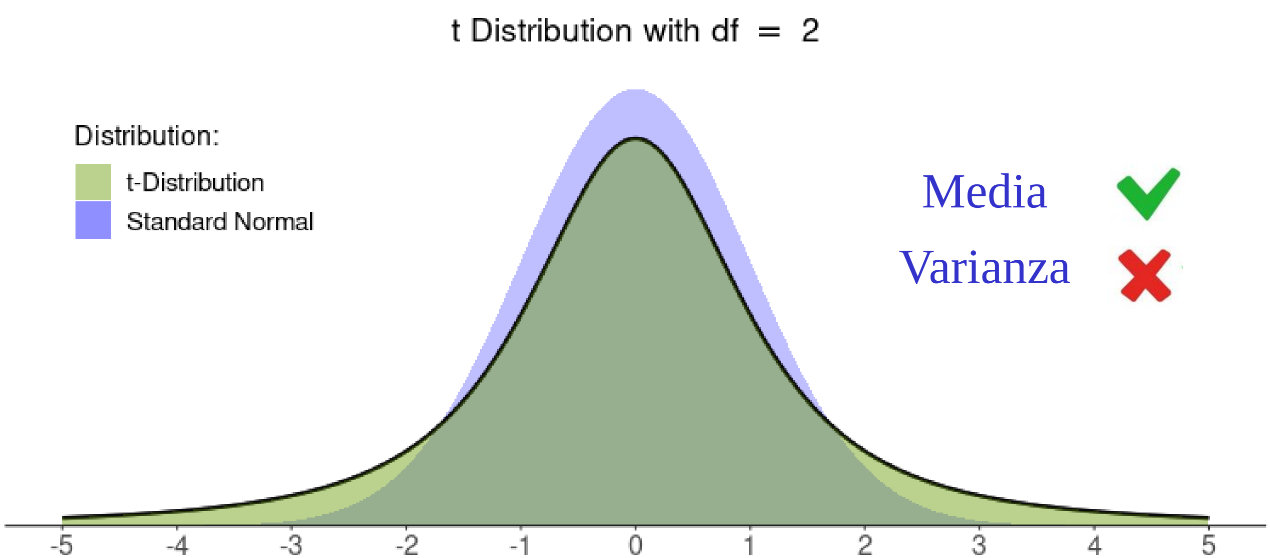
La v.c. t **di Student** ha distribuzione **limite** Normale standardizzata

Per g **sufficientemente grande** (> 30)		La v.c. t **di Student** (di parametro g) ha distribuzione **asintotica** Normale standardizzata
$T \sim N(0,1)$	In modo **approssimato**	



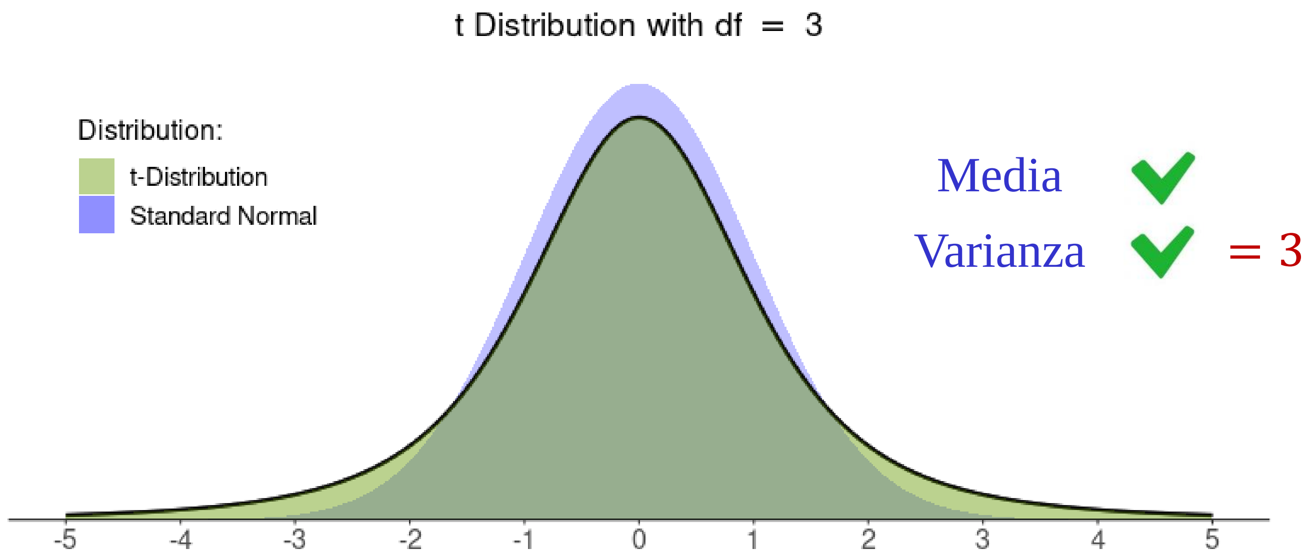
Media ✗

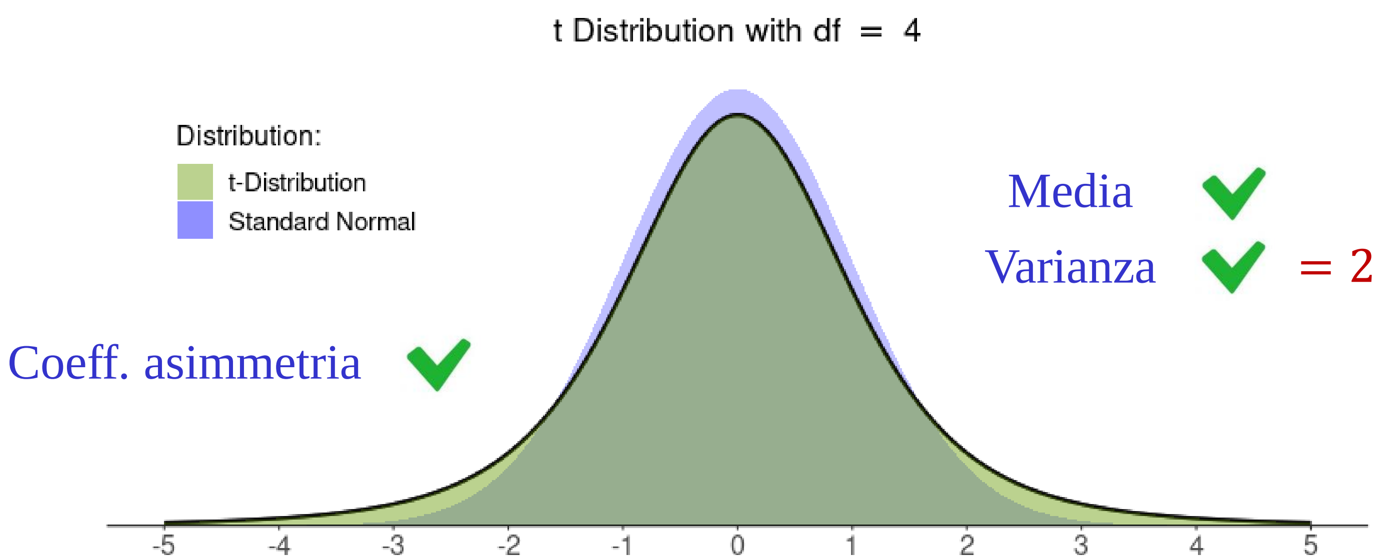
Varianza ✗

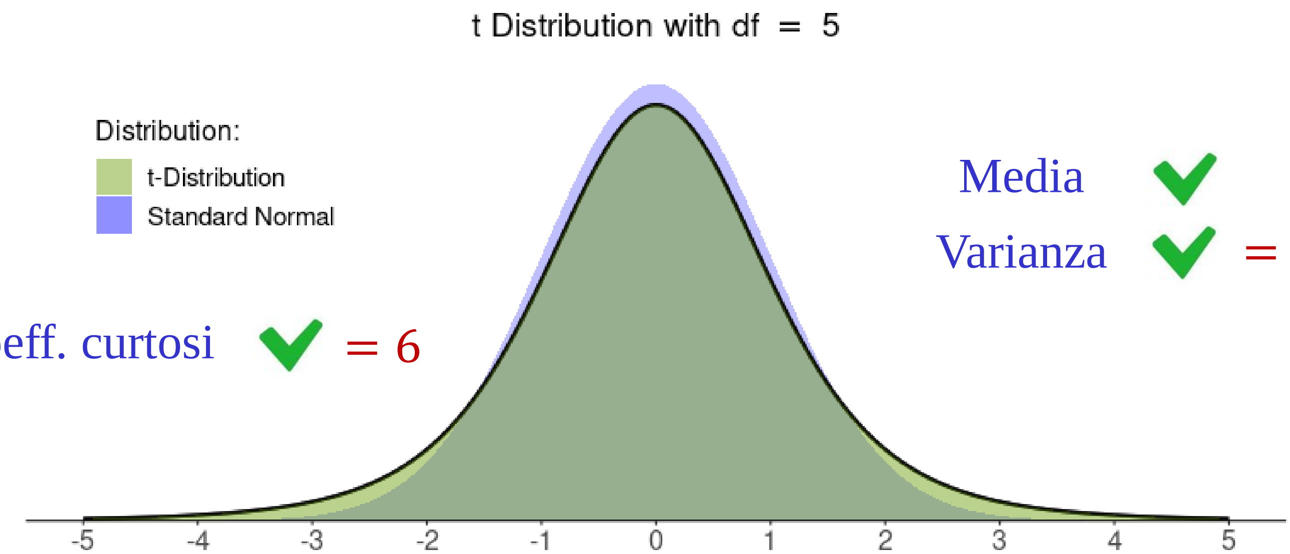


Media ✓

Varianza ✗


t Distribution with df = 3
Distribution:
t-Distribution
Standard Normal
Media
Varianza = 3
-5 -4 -3 -2 -1 0 1 2 3 4 5


t Distribution with df = 4
Distribution:
t-Distribution
Standard Normal
Media
Varianza = 2
Coeff. asimmetria
-5 -4 -3 -2 -1 0 1 2 3 4 5

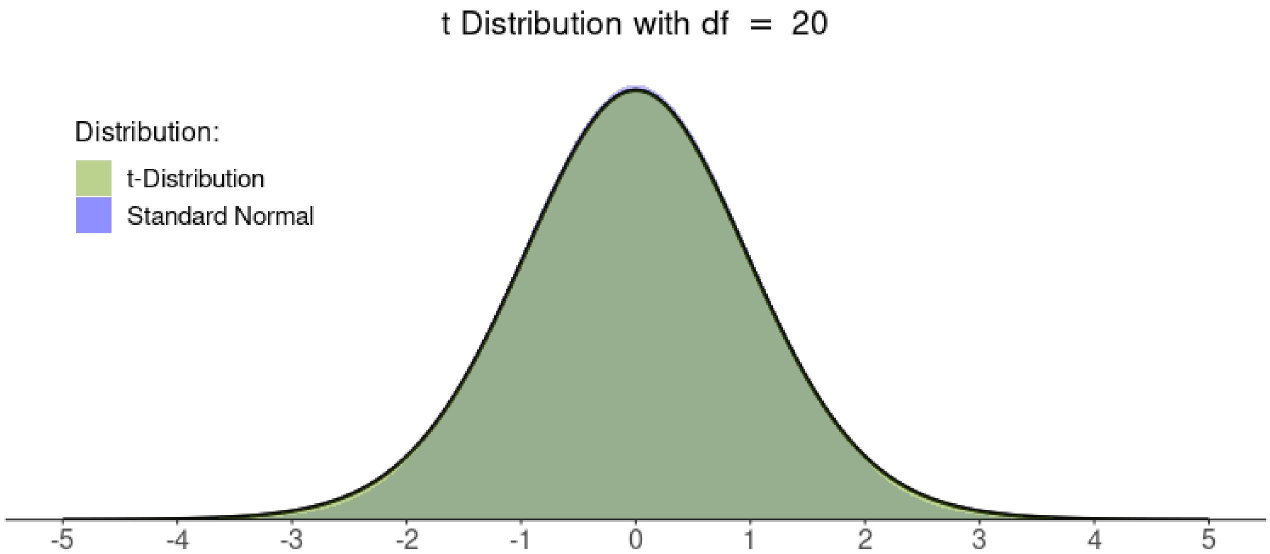


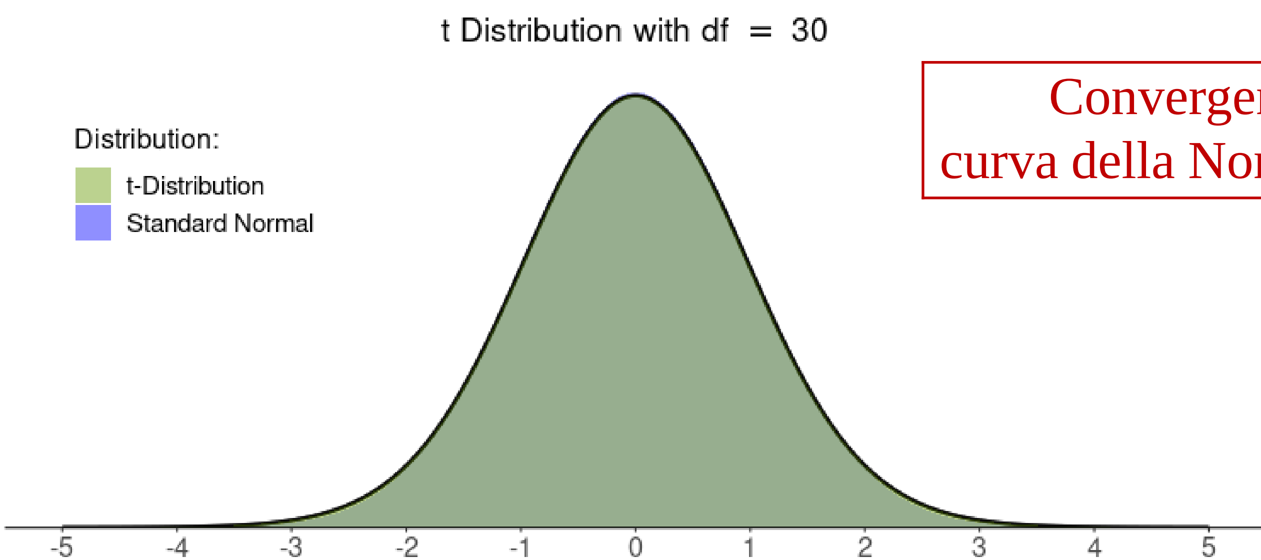
Media ✓

Varianza ✓ $= 1.7$

Coeff. curtosi ✓ $= 6$

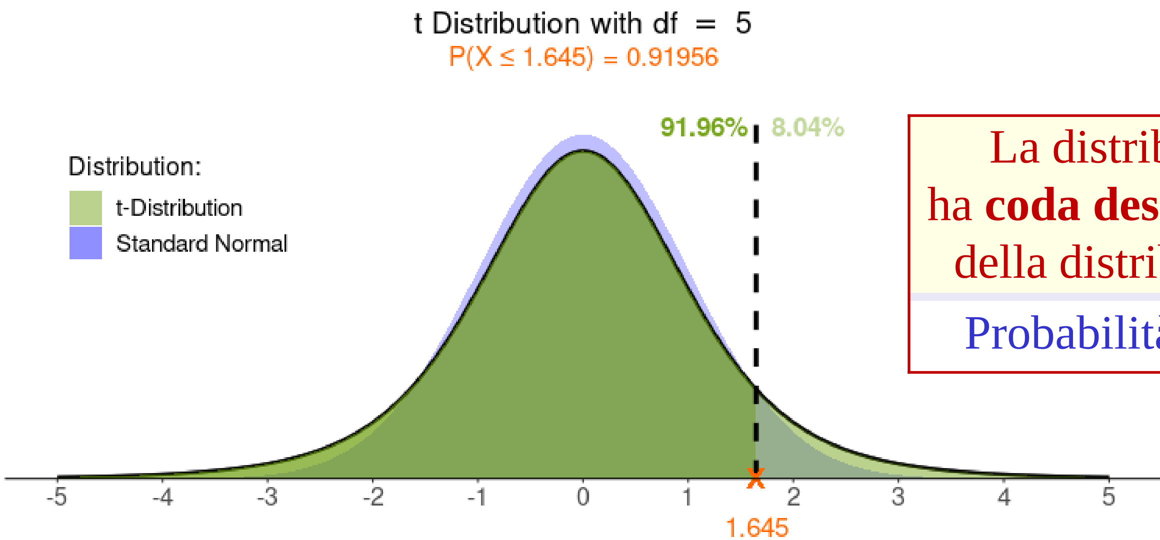
Coefficiente di curtosi	$g \geq 5$	$0 < E(Z^4) - 3 = \dfrac{6}{g-4} \leq 6$	$\xrightarrow[g \to \infty]{} 0$
Per ogni valore di g	Distribuzione t di Student **leptocurtica** (ipernormale)		
Forma **più appuntita** rispetto a quella della **Normale standardizzata**			
La distribuzione t di Student ha **code più pesanti**			
Maggiore probabilità ad intervalli distanti dalla parte centrale della distribuzione			



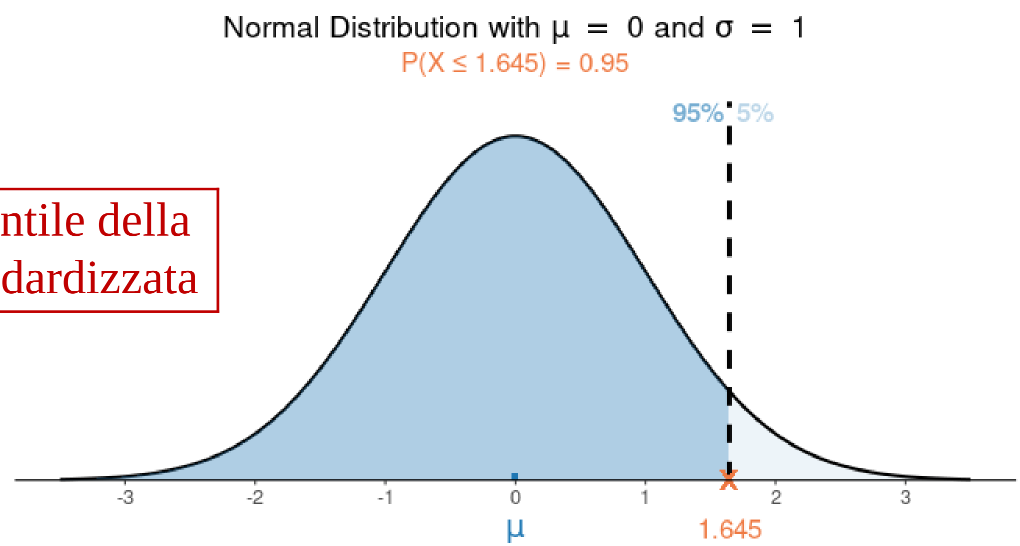


Convergenza **rapida** alla
curva della Normale standardizzata

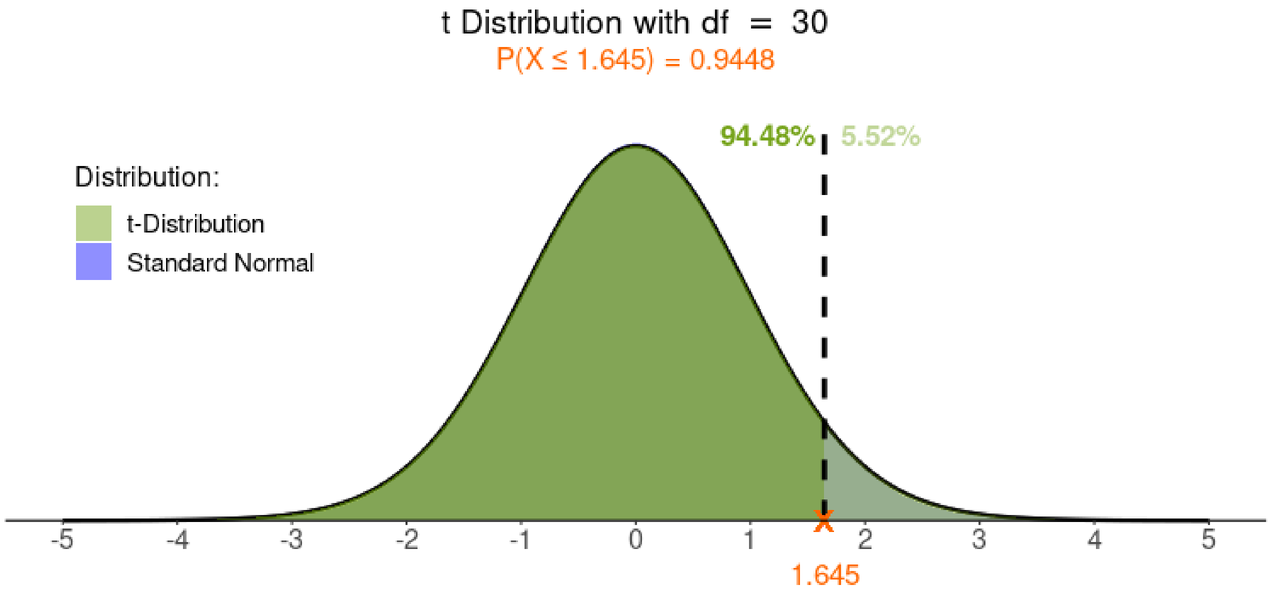
Funzione di ripartizione (e **quantili** da essa derivati) di una v.c. $T \sim t(g)$	
Non calcolabili in modo esplicito	Calcolo basato su *integrazione numerica*
Quantili di livello p **tabulati** per alcuni valori rilevanti di p (al variare di g)	

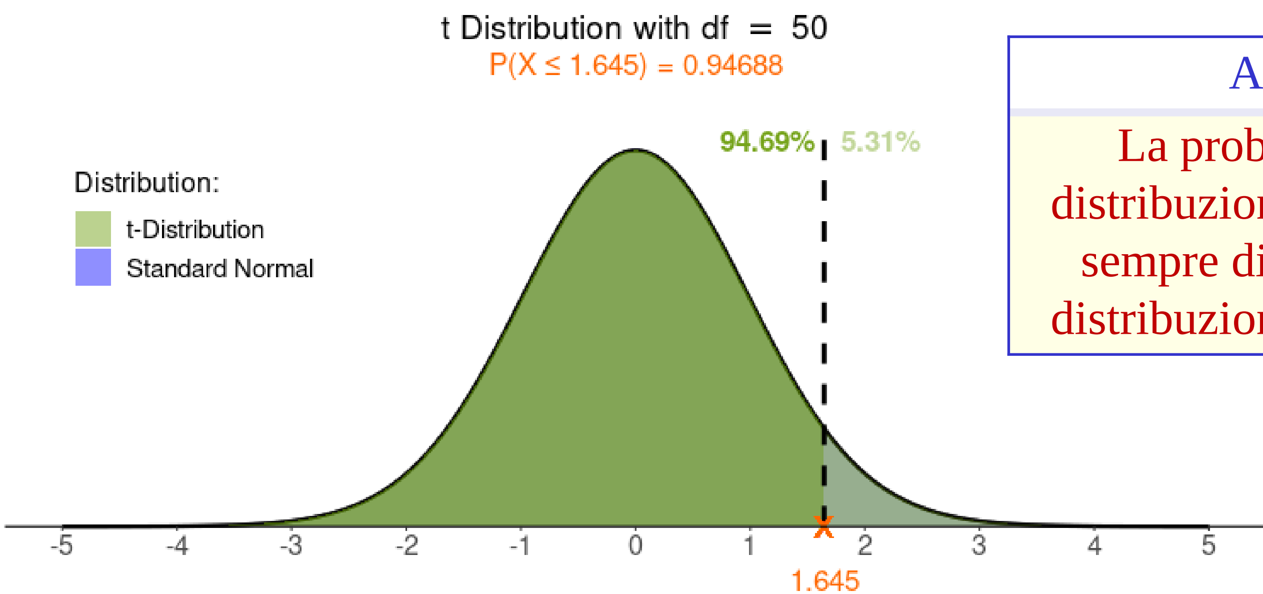


La distribuzione t di Student (con 5 gl) ha **coda destra** (valori > 1.645) **più pesante** della distribuzione Normale standardizzata	
Probabilità della coda	$0.0804 > 0.05$



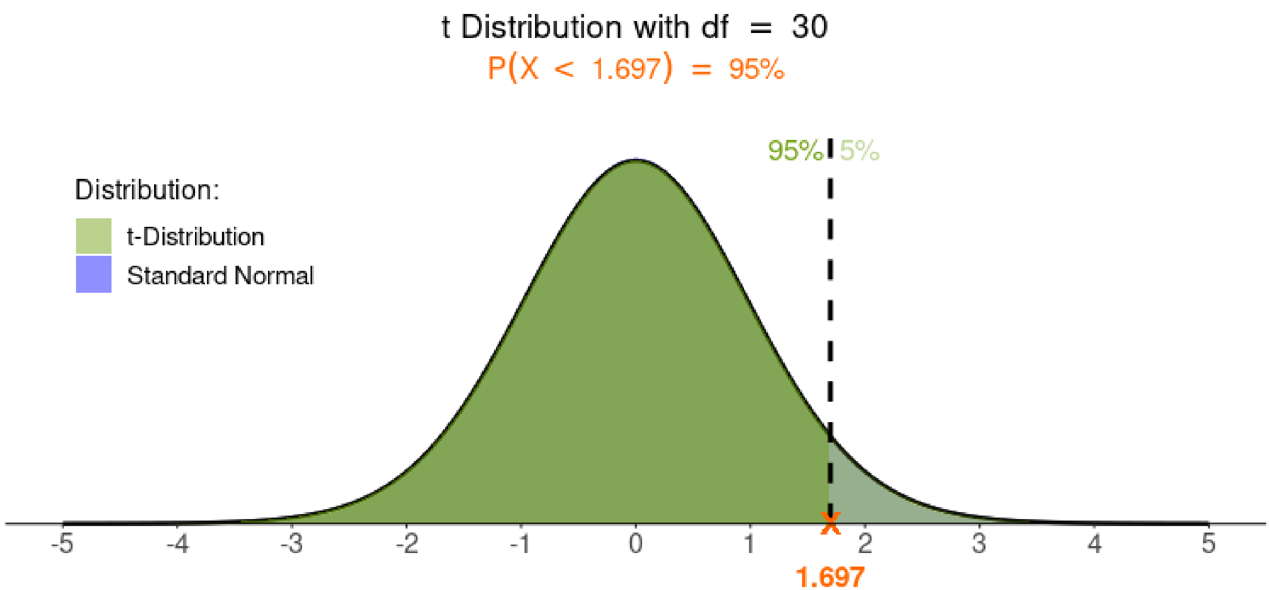
$\mathbf{1.645}$ è il 95-esimo percentile della distribuzione Normale standardizzata

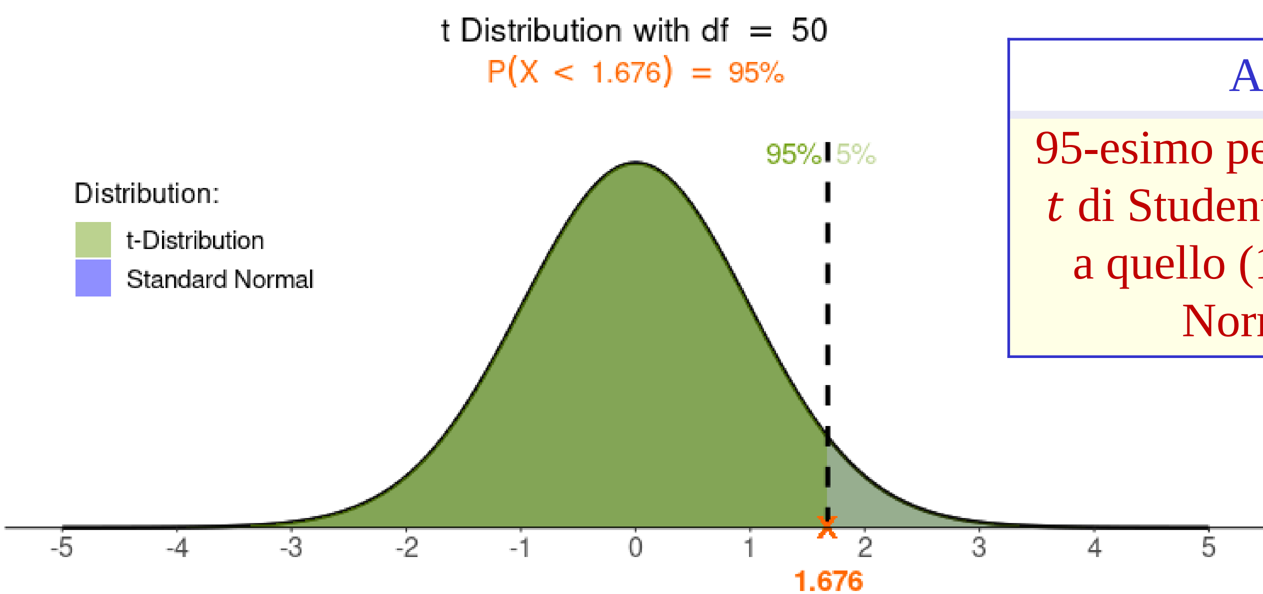




All'aumentare di *g*

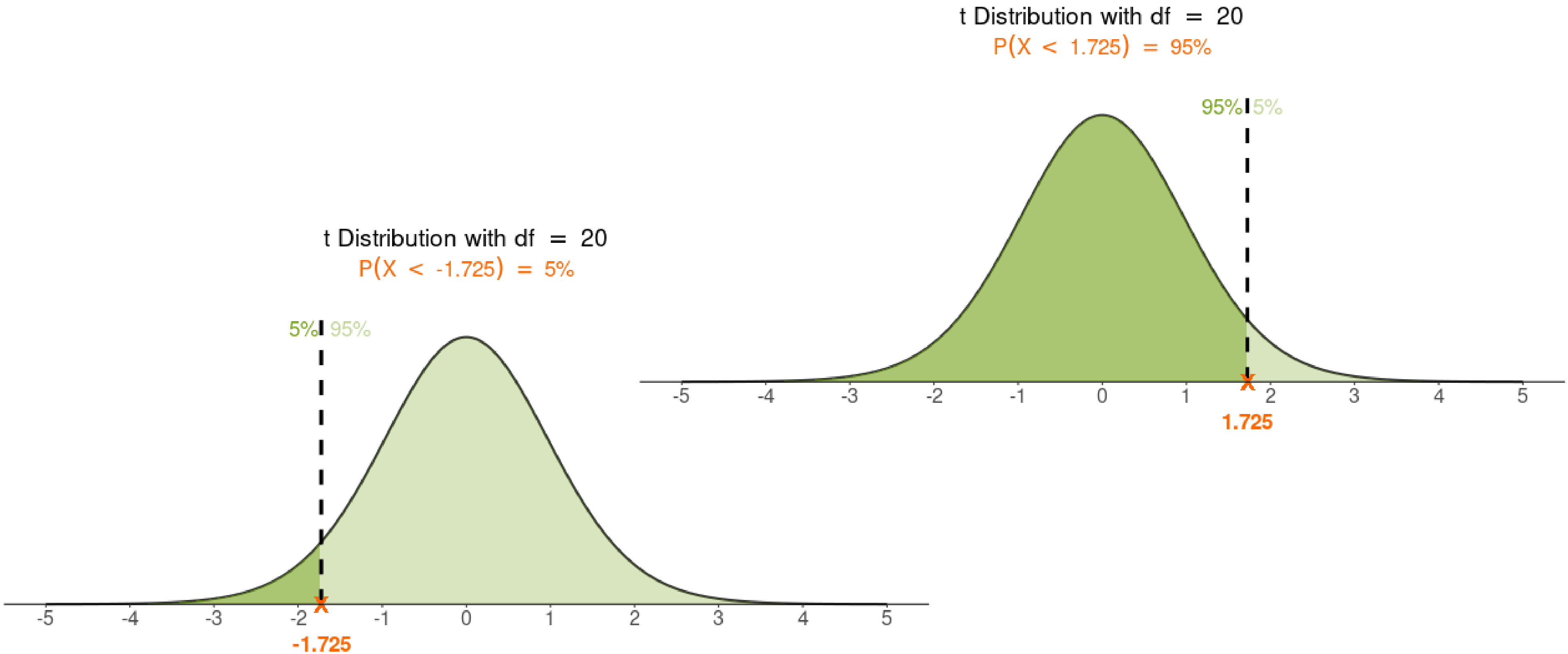
La probabilità della coda della distribuzione *t* di Student **si avvicina** sempre di più a quella (0.05) della distribuzione Normale standardizzata





All'aumentare di *g*

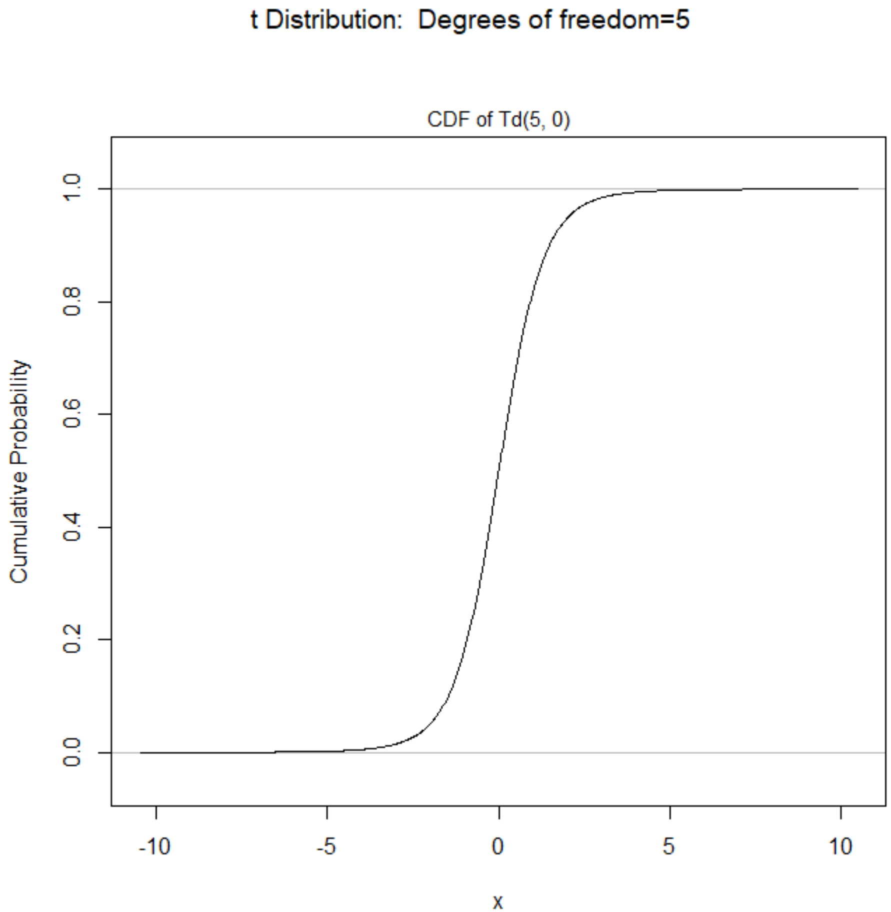
95-esimo percentile della distribuzione *t* di Student **si avvicina** sempre di più a quello (1.645) della distribuzione Normale standardizzata

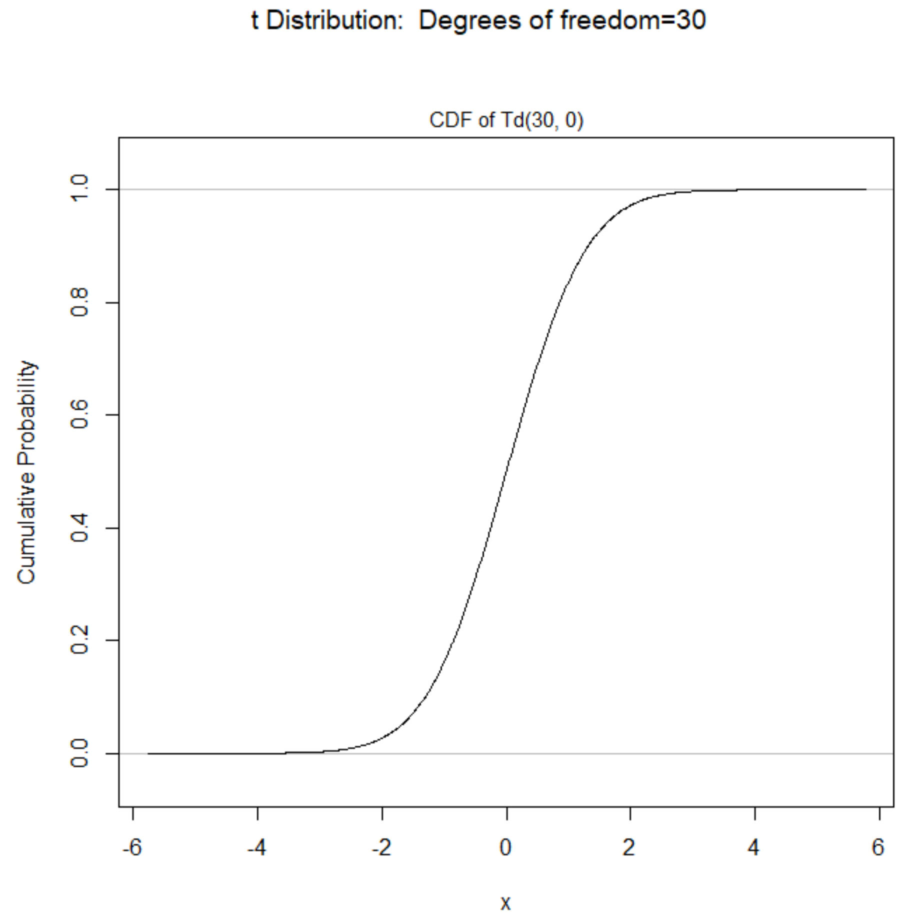


Per ogni valore **positivo** $t > 0$	$F(-t) = 1 - F(t)$

A probabilità **complementari** corrispondono quantili **opposti**

Quantili di livello p **tabulati** solo per valori di $p > 0.5$ (alcuni)





All'aumentare di *g*

La **curva di ripartizione** *t* di Student **somiglia** sempre di più alla **ogiva** della Normale standardizzata

Modelli di probabilità per variabili casuali continue — Quadro riassuntivo

	Uniforme		**Normale**		***t* di Student**	
Notazione	$X \sim U(a,b)$		$X \sim N(\mu,\sigma^2)$		$T \sim t(g)$	
Funzione di densità	$\frac{1}{b-a}$		$\frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right\}$			
Supporto	$[a,b]$		$\mathbb{R}$		$\mathbb{R}$	
Parametri	a,b		μ,σ^2		g	
Spazio parametrico	$\mathbb{R}^2$		$\{(\mu,\sigma^2)\colon\ \mu \in \mathbb{R};\ \sigma^2 \in \mathbb{R}_0^+\}$		$\mathbb{Z}^+ = \{1,2,\cdots,\infty\}$	
Media e varianza	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	μ	σ^2	0	$\frac{g}{g-2}$
Funzione di ripartizione	$\frac{x-a}{b-a}$		Integrazione numerica		Integrazione numerica	
Impiego principale	**Giochi / Teoria della simulazione**				**v.c. campionaria**	