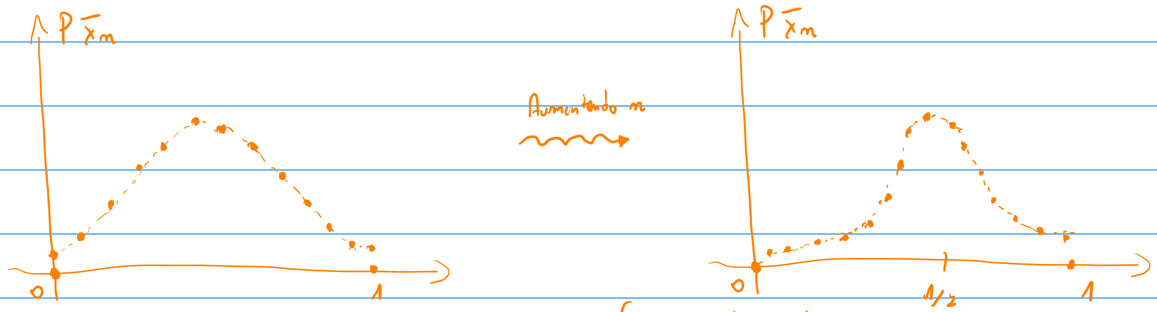
TEOREMI LIMITE

E.g. lanci di monete non truccate

① $X_1 \dots X_n \dots$ Bernoulli $(1/2)$ $\qquad \bar{X}_n = \dfrac{X_1 + \dots X_n}{n}$ Ha esiti possibili $0, 1/n, 2/n, \dots, 1$

$(X_1 + \dots + X_n$ é Binomiale $(n, \frac{1}{2}))$

$P_{X_1 + \dots + X_n}(k) = \binom{n}{k}\left(\frac{1}{2}\right)^n$



$$\forall \varepsilon \geq 0 \quad \lim_{n \to \infty} \mathbb{P}\left(|\bar{X}_n - \tfrac{1}{2}| > \varepsilon\right) = 0$$

Definizione: Siano $X_1, X_2, \dots$ una successione di VARIABILI ALEATORIE (su Ω, $\mathbb{P}$) con funzione di ripartizione F_{X_n} e s.a $X: \Omega \to \mathbb{R}$ con funzione di ripartizione F_X e con densitá, allora diciamo che X_n converge per $n \to \infty$ a X in legge (in distribuzione) se $\forall t \in \mathbb{R}$ $\lim_{n \to \infty} F_{X_n}(t) = F_X(t)$

Teorema [Centrale del Limite]: se $X_1, X_2, \dots$ sono una successione di variabili aleatorie INDIPENDENTI e IDENTICAMENTE DISTRIBUITE con MOMENTO SECONDO FINITO ($\exists$ finiti $\mathbb{E}[X_n] = \mu$ $VAR(X_n) = \sigma^2$) allora le variabili $Z_n = \dfrac{X_1 + \dots + X_n - \mu n}{\sqrt{n}\,\sigma}$ convergono in legge ad una GAUSSIANA STANDARD $N(0,1)$

Nota: $\dfrac{X_1 + \dots + X_n - \mu n}{\sqrt{n}\,\sigma} = \sqrt{n}\,\dfrac{\bar{X}_n - \mu}{\sigma}$

$\downarrow$ ovvero

Per ogni intervallo $[a,b] \subseteq \mathbb{R}$ $\mathbb{P}(a \leq Z_n \leq b) = \mathbb{P}(a \leq Z \leq b) = \Phi(b) - \Phi(a)$

dove $Z = N(0,1)$ Gaussiana Standard

Nell'esempio ① precedente $\mu = 1/2$, $\sigma^2 = \frac{1}{4}$

$$\mathbb{P}\left(a \leq \frac{X_1 + \dots + X_n - \frac{n}{2}}{\sqrt{n}/2} \leq b\right) \approx \Phi(b) - \Phi(a)$$

E.g. Continuando l'esempio ①, se lanciamo 1000 monete, qual è la probabilità di ottenere almeno 900 teste?

$\#$ di teste $X_1+\dots+X_n$ è Binomiale$(1000, 1/2)$

$$\mathbb{P}(X_1+\dots+X_n \geq 900) = \sum_{k=900}^{1000} \binom{1000}{n}\left(\frac{1}{2}\right)^{1000} \longrightarrow$$

Nessuno dei due calcoli è facile da fare ma un tempo per Φ c'erano delle tabelle di valori

$\uparrow$ EFFICIENZA

$\downarrow$ Per il TEOREMA CENTRALE del LIMITE

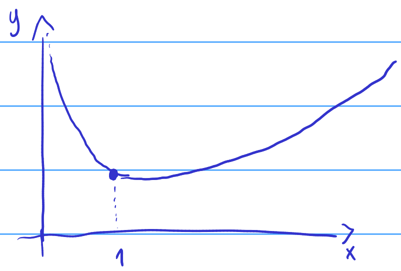
$$\mathbb{P}(X_1+\dots+X_n \geq 900) = \mathbb{P}\left(\frac{X_1+\dots+X_n - 1500}{1/2\cdot\sqrt{1000}} \geq \frac{900-500}{1/2\cdot\sqrt{1000}}\right) \approx 1-\Phi\left(\frac{400}{1/2\,\sqrt{1000}}\right)$$

$\downarrow$ GRANDEZZA di n

Questo perché n è grande
Per i nostri scopi almeno $n \geq 80$

VARIABILI ASSOCIATE e GAUSSIANE INDIPENDENTI

Definizione [Gamma di Eulero] La funzione Gamma di Eulero è una funzione definita su $\Gamma:(0,\infty)\to(0,\infty)$ e vale $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt$



$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt = \left[-t^{x-1}e^{-t}\right]_0^\infty + \int_0^\infty t^{x-2}e^{-t}\,dt\,(x-1) = (x-1)\Gamma(x-1)$$

$$\Downarrow$$

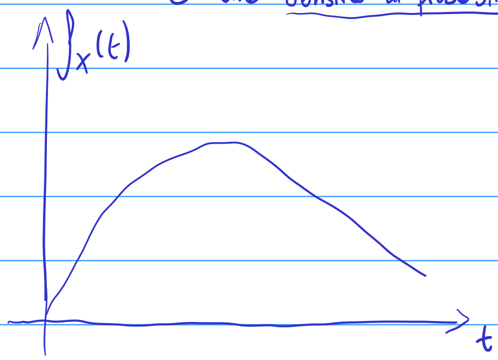
$\Gamma(n) = (n-1)!$, quindi estende la funzione $n \longmapsto n!$ a $\mathbb{R}$

Definizione: una variabile aleatoria X ha densità Γ di parametri $r, \lambda > 0$ se

$$f_X(t) = \begin{cases} \frac{1}{\Gamma(r)}\lambda^r t^{r-1}e^{-\lambda t} & t>0 \\ 0 & t\leq 0 \end{cases}$$

È una densità di probabilità perché $\int_0^\infty f_X(t)\,dt = \frac{\lambda^r}{\Gamma(r)}\int_0^\infty t^{r-1}e^{-\lambda t}\,dt =$

$$\overset{s=\lambda t}{=} \frac{\lambda^r}{\Gamma(r)}\int_0^\infty \frac{s^{t-1}}{\lambda^{r-1}}e^{-s}\,\frac{ds}{\lambda} = 1$$



$r \to$ Parametro di FORMA
$\lambda \to$ Parametro di DECADIMENTO
$\Big\} \to$ Si scrive come $\Gamma(r,\lambda)$

Nota Una tale variabile ha tutti i MOMENTI definiti

$$\mathbb{E}[x^n] = \frac{1}{\Gamma(r)} \int_0^\infty t^n \lambda^r t^{r-1} e^{-\lambda t} dt = \frac{\lambda^{-n} \lambda^{n+r}}{\Gamma(r)} \int_0^\infty t^{n+r-1} e^{-\lambda t} dt =$$

$$\overset{s=\lambda t}{=} \frac{1}{\Gamma(r)\lambda^n} \cdot \lambda^{n+r} \int_0^\infty \frac{s^{n+r-1}}{\lambda^{n+r-1}} e^{-s} \frac{ds}{\lambda} = \frac{\Gamma(n+r)}{\Gamma(r)\lambda^n}$$

Lemma ②: se X, Y hanno densitá rispettivamente $\Gamma(r,\lambda)$ e $\Gamma(s,\lambda)$ e sono INDIPENDENTI allora $X+Y$ ha densitá $\Gamma(r+s,\lambda)$

(Si dimostra con la formula della CONVOLUZIONE $f_Z(t) = \int_{-\infty}^{\infty} f_X(t-s) f_Y(s)\, ds$)

Lemma: se $X_1, \dots, X_n$ sono Gaussiane $N(0,1)$ INDIPENDENTI, allora la variabile $Y = X_1^2 + \dots + X_n^2$ ha DENSITÁ $\Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$
(in particolare $X_i^2 = \Gamma\left(\frac{1}{2}, \frac{1}{2}\right)$)

$\downarrow$

La densitá $\Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$ si dice anche $\chi^2(n)$ "chi-quadro" con n GRADO di LIBERTÁ

Dimostrazione: se X_i^2 é $\Gamma(1/2, 1/2)$ allora per il lemma ② $X_1^2 + \dots + X_n^2$ é $\Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$
resta da verificare che $f_{X_i^2}(t) = \frac{1}{\Gamma\left(\frac{1}{2}\right)} \sqrt{1/2}\; t^{1/2-1} e^{-t/2} \quad t > 0$

Non possiamo usare il CAMBIO di VARIABILI $\quad t \geq 0$
$X \sim N(0,1) \quad W = X^2 = h(X) \quad h(x) = x^2$
perché h non é BIGETTIVA su $\mathbb{R}$

$\downarrow$

$$F_W(t) = \mathbb{P}(W \leq t) = \mathbb{P}(X^2 \leq t) = \mathbb{P}(-\sqrt{t} \leq X \leq \sqrt{t}) = \Phi(\sqrt{t}) - \Phi(-\sqrt{t}) = 2\Phi(\sqrt{t}) - 1$$

$$\Phi(t) = \int_{-\infty}^{t} \frac{1}{2\sqrt{t}} e^{-s^2/2} ds \qquad f_W(t) = \frac{dx}{dt} F_W(t) = 2\frac{dx}{dt}\Phi(\sqrt{t}) = \frac{1}{\sqrt{2t}} e^{-t/2}$$

Oss. La densitá $\Gamma(1,\lambda)$ é la densitá ESPONENZIALE di parametro λ

Osservazioni:

① Se X, Y indipendenti hanno densità $\chi^2(n)$ e $\chi^2(m)$
allora $X+Y$ ha densità $\chi^2(n+m)$

② Se $X_1, \dots, X_n$ sono $N(0,1)$ Gaussiane indipendenti
Vale la Legge dei Grandi Numeri
Vale il Teorema Centrale del Limite

Esercizio

L'altezza X di un individuo a caso ha distribuzione Gaussiana $N(177, 10{,}6^2)$

- Qual è la probabilità che sia alto almeno 180?

$$\mathbb{P}(X > 180) = \mathbb{P}\left(\frac{X-177}{10.6} > \frac{180-177}{10.6}\right) = 1 - \Phi\left(\frac{3}{10.6}\right)$$

Se X è $N(m, \sigma^2)$
allora $\frac{X-m}{\sigma}$ è $N(0,1)$

- Qual è la probabilità che la media aritmetica $\frac{X+Y}{2}$ dell'altezza di due individui a caso (X, Y indip.) sia almeno 180?

Se $X, Y \sim N(m, \sigma^2)$ allora $X+Y$ è $N(2m, 2\sigma^2)$

$$\mathbb{P}\left(\frac{X+Y}{2} > 180\right) = \mathbb{P}\left(\frac{X+Y-2m}{\sqrt{2}\,\sigma} > \frac{2 \cdot 180 - 2m}{\sqrt{2}\,\sigma}\right) = 1 - \Phi\left(\frac{2 \cdot 180 - 2 \cdot 177}{\sqrt{2} \cdot 10.6}\right)$$

- Presi 100 individui $X_1, \dots, X_n$ a caso (v.a. indipendenti con stessa distribuzione), qual è la probabilità che la media aritmetica delle loro altezze $\frac{X_1 + \dots + X_n}{100}$ sia almeno 180?

$$\mathbb{P}\left(\frac{X_1 + \dots + X_n}{100} > 180\right) = 1 - \Phi\left(\frac{100 \cdot 180 - 100 \cdot 177}{\sqrt{100} \cdot 10.6}\right)$$

$X_1 + \dots + X_n$ ha densità $N(100m, 100\sigma^2)$

- Se io non sapessi che $X_1, \dots, X_n$ sono Gaussiane ma solo che sono i.i.d. con $\mathbb{E}[X_i] = 177$ $VAR(X_i) = 10.6$, cosa posso dire?

Per il TCL sappiamo che $\sqrt{n}\,\frac{\overline{X}_n - m}{\sigma}$ converge a $N(0,1)$ in legge,

$$\mathbb{P}\left(\frac{X_1 + \dots + X_n}{n} > 180\right) = \mathbb{P}\left(\sqrt{n}\,\frac{\overline{X}_n - m}{\sigma} > \sqrt{n}\,\frac{180 - m}{\sigma}\right) \approx 1 - \Phi\left(\sqrt{n}\,\frac{180 - m}{\sigma}\right)$$