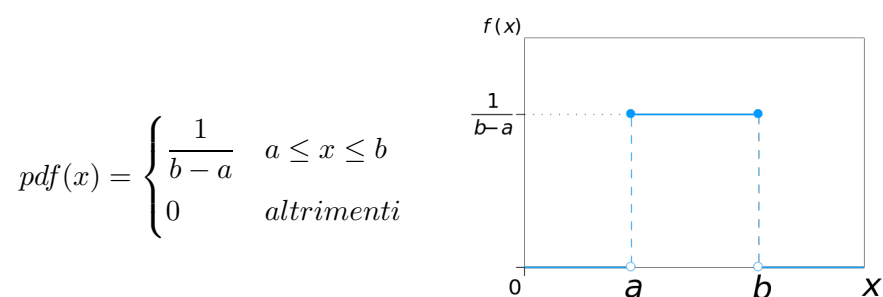


CHAPTER 2

Modelli di probabilità

2.1 La distribuzione Uniforme

La distribuzione di probabilità uniforme per una variabile aleatoria x continua è definita come:



Tale distribuzione definisce una distribuzione di probabilità uguale per tutti le variabili aleatorie $x \in \Omega$.

2.1.1 Proprietà della pdf

Il valore atteso della distribuzione uniforme e la sua varianza sono date da:

$$E[x] = \int_a^b \frac{x}{b-a} dx = \frac{b+a}{2} \quad (2.1)$$

$$V[x] = \int_a^b \frac{1}{b-a} \left(x - \frac{b+a}{2}\right)^2 dx = \frac{(b-a)^2}{12} \quad (2.2)$$

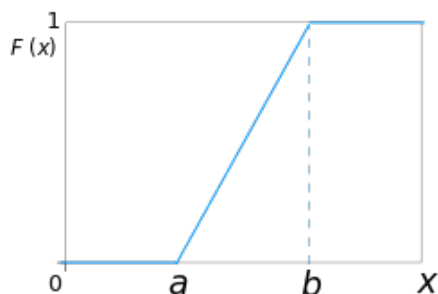
Per una distribuzione di questo tipo si ha che la simmetria e la kurtosi rispettivamente hanno valore nullo.

Inoltre tale pdf gode della proprietà di **riproduttività** ovvero date due variabili casuali x e y che seguono la pdf uniforme si ha che la nuova variabile aleatoria $z = x + y$ segue una distribuzione di probabilità pdf(z) con media $\mu_z = \mu_x + \mu_y$ e varianza $\sigma_z^2 = \sigma_x^2 + \sigma_y^2$.

Un'altra proprietà importante della pdf è data dal fatto che se si ha una variabile aleatoria x che segue una distribuzione uniforme, la variabile casuale $y = f(x)$ seguirà anch'essa la medesima distribuzione di probabilità.

Poichè la pdf possiede una forma analitica è possibile definirne la cdf:

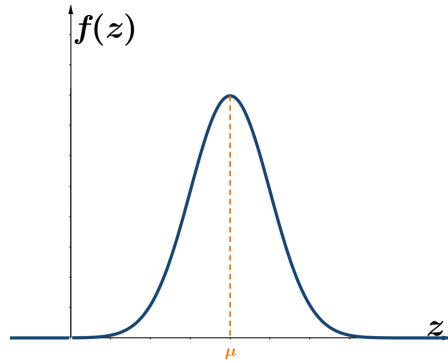
$$cdf(x) = \int_a^x \frac{dx}{b-a} = \frac{1}{b-a}(x-a)$$



2.1.2 La distribuzione Gaussiana (o Normale)

Data una variabile aleatoria $x \in \Omega$ continua la distribuzione di probabilità Gaussiana è definita come:

$$G(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x - \mu)^2}{2\sigma^2}} \quad (2.3)$$



Per una distribuzione Gaussiana si ha che la media, la moda e la mediana coincidono essendo una distribuzione simmetrica ovvero $\gamma_1, \gamma_2 = 0$. Media, Varianza e deviazione standard coincidono con quelle definite al capitolo precedente. Gode della proprietà di **riproduttività** definita nella sezione della distribuzione uniforme.

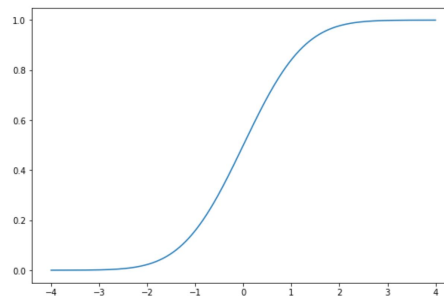
Una caso speciale di Gaussiana è quella **standardizzata** che è espressa come:

$$G(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

2.1.3 Distribuzione cumulativa di una Gaussiana

La pdf Gaussiana non ammette primitiva e la sua cdf viene calcolata numericamente e prende il nome di **funzione d'errore**, abbreviata con $\text{erf}(x)$.

$$\text{erf}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$



2.1.4 Teorema Centrale del Limite

La distribuzione Gaussiana è particolarmente importante poichè per essa vale un importante risultato.

Teorema

Si ipotizzi di avere N variabile aleatorie $\{x_i\}_i^N$ indipendenti tra loro e ciascuna di esse segue una distribuzione di probabilità $f_i(x)$, e definiamo una nuova variabile aleatoria $s = \sum_{i=1}^N x_i$ che segue una distribuzione pdf(s), allora si avrà che al crescere del numero di variabili:

$$\lim_{N \rightarrow \infty} pdf(s) \rightarrow G(s, \mu_s, \sigma_s)$$

$$\text{dove } \mu_s = \sum_{i=1}^N \mu_i \text{ e } \sigma_s^2 = \sum_{i=1}^N \sigma_i^2.$$

Corollario della media

Date N variabili aleatorie $\{x_i\}_i^N$ indipendenti ed identicamente distribuite, ovvero che per ogni variabile si ha la stessa distribuzione di probabilità $f(x)$ e ipotizzando che $\forall i$ si ha che μ_i e σ_i^2 siano finite, definita una nuova variabile aleatoria $\bar{x} = \frac{\sum_{i=1}^N x_i}{N}$ che segue una pdf $f(\bar{x})$, si ha che al crescere del numero di variabili:

$$\lim_{N \rightarrow \infty} f(\bar{x}) \rightarrow G(\bar{x}, \mu, \frac{\sigma}{\sqrt{N}})$$

dunque si ha che la pdf delle medie campionarie di un campione che cresce converge alla distribuzione Gaussiana.

Conseguenze del TCL

Quando effettuiamo una misura in un esperimento in generale possiamo definirla come la somma del valore vero x_0 e di un errore ϵ che è una variabile aleatoria.

$$x = x_0 + \epsilon$$

la misura così ottenuta segue una pdf(x), il che equivale a dire che segue la pdf(ϵ), quando si effettua la misurazione di una grandezza possiamo pensare ad ϵ che l'incertezza di stima come il contributo di molteplici errori, dunque al crescere di tale numero per il TCL si avrà che $\text{pdf}(\epsilon) \rightarrow G(\epsilon, \mu, \frac{\sigma}{\sqrt{N}})$.

2.2 Distribuzione di probabilità Log-normale

Preso una variabile aleatoria $x \in \Omega$ che segue una distribuzione Gaussiana, costruiamo una nuova variabile aleatoria y legata alla precedente dalla relazione $y = e^x$, in questo modo si definisce la distribuzione log-normale applicando un cambio di variabile alla $G(x)$ di partenza, il metodo con cui si effettua un cambio di variabile verrà presentato nei capitoli successivi.

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

2.2.1 Proprietà della log-normale

Il valore di aspettazione e varianza della log-normale sono dati da:

$$E[x] = e^{(\mu + \frac{1}{2}\sigma^2)} \quad (2.4)$$

$$V[x] = e^{(2\mu + \sigma^2)} e^{(\sigma^2 - 1)} \quad (2.5)$$

2.3 Distribuzione di Probabilità di Bernoulli

Sia $x \in \Omega = \{0, 1\}$ una variabile aleatoria discreta, a cui è possibile associare una probabilità $p \in [0, 1]$. Definiremo che l'evento x si è realizzato con successo quando $P(x = 1) = p$, mentre l'insuccesso con $P(x = 0) = (1-p)$. Dunque la pdf(x) è data da:

$$f(x, p) = p^x (1-p)^{1-x}$$

2.3.1 Proprietà della distribuzione Bernoulliana

Per una distribuzione di Bernoulli media e varianza sono date:

$$E[x] = \sum_{\{0,1\}} x f(x, p) = p \quad (2.6)$$

$$V[x] = \sum_{\{0,1\}} (x - p)^2 f(x, p) = p(1 - p) \quad (2.7)$$

2.4 Distribuzione di Probabilità Binomiale

Si consideri un insieme di N osservazioni indipendenti, che seguono una distribuzione di Bernoulli. Contiamo il numero di volte k in cui l'evento ha successo rispetto al numero di tentativi effettuati, tale insieme è il nostro campione. Se si ripete l'esperimento effettuando sempre N tentativi, il numero di successi n sarà distribuita secondo la distribuzione binomiale:

$$B(k|N, p) = \binom{N}{k} p^k (1 - p)^{N-k}$$

notare che a differenza delle distribuzioni viste fino ad ora, la distribuzione binomiale in realtà rappresenta la probabilità stessa di una variabile aleatoria k , come anche la Bernoulliana.

2.4.1 Proprietà della distribuzione binomiale

Per una distribuzione binomiale il valore di aspettazione e la varianza sono definite da :

$$E[k] = \sum_k k \binom{N}{k} p^k (1 - p)^{N-k} = Np \quad (2.8)$$

$$V[k] = E[k^2] - E[k]^2 = Np(1 - p) \quad (2.9)$$

la simmetria è data da:

$$\gamma_1 = \frac{1 - 2p}{\sqrt{Np(1 - p)}} \quad (2.10)$$

dove $\gamma_1 \rightarrow 0$ per $N \rightarrow \infty$ o per $p = \frac{1}{2}$.

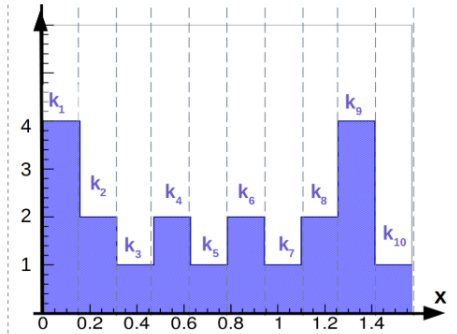
la kurtosi invece si esprime come:

$$\gamma_2 = \frac{1 - 6p}{Np(1 - p)} \quad (2.11)$$

dove $\gamma_2 \rightarrow 0$ per $N \rightarrow \infty$ e $p = \frac{1}{6}$. Inoltre la distribuzione binomiale gode della proprietà di **riproduttività**

2.4.2 La densità di probabilità di un'istogramma

Consideriamo di avere una variabile aleatoria x di cui effettuiamo N misurazioni, ciascuna misurazione segue una pdf(x).



L'operazione è equivalente a campionare N volte una pdf(x). Posto che le $\{x_i\}_i^N$ misure siano il campione di riferimento, la statistica si pone obbiettivo quello di ricostruire partendo dal campione le caratteristiche di una pdf(x).

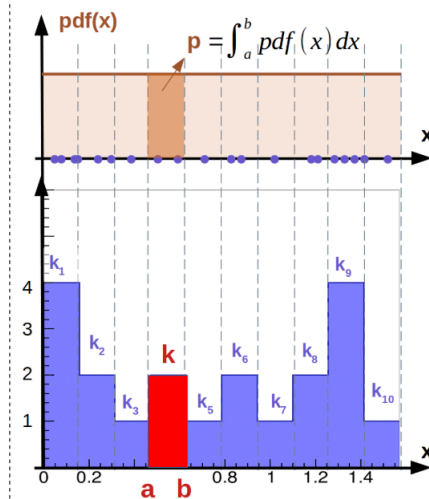
Per farlo si procede a costruire un istogramma e a contare il numero di eventi k_i che cadono all'interno di

un bin e rappresentare tale numero con una barra. Il numero di misure k_i che cade all'intero di un bin è una variabile aleatoria.

Di conseguenza studiando il comportamento di un singolo bin $[a,b)$, il numero di misure k_i è descritto da una distribuzione di probabilità, poichè una misura cade o non cade all'interno dell'intervallo la pdf(k_i) è quella di Bernoulli.

Reiterando l'esperimento sullo stesso bin il numero di successi k_i seguirà una pdf(k) di una binomiale $B(k|N, p)$ dove la probabilità è $p = \int_a^b pdf(x) dx$ sull'intervallo del bin associato.

Di conseguenza un conteggio di misure è soggetto a fluttuazioni statistiche, e al crescere del numero delle



misure l'istogramma assumerà la distribuzione di probabilità pdf(x) della variabile aleatoria X misurata.

Se ipotizziamo che X abbia una distribuzione uniforme, valore di aspettazione e varianza di un conteggio coincideranno con quella della distribuzione Binomiale.

I metodi statistici che consentono di utilizzare il metodo dell'istogramma richiedono che i conteggi k_i siano variabili aleatorie indipendenti, si hanno due strategie affinché questo sia realizzabile:

- Se la dimensione dei bin è piccola e la probabilità è sufficientemente piccola allora la covarianza è pressochè nulla. Di conseguenza le variabili k_i sono indipendenti tra loro e dunque ciascuna distribuita come una binomiale.
- Se il numero di eventi N è molto grande e la probabilità p_i è piccola la distribuzione di probabilità dei conteggi è approssimabile a quella di una distribuzione Poissoniana.

2.5 Distribuzione Multinomiale

Considerando quanto discusso nella sezione precedente per ogni singolo bin che si è costruito nell'istogramma, si definisce una probabilità multinomiale:

$$Multi(k_1, k_2, \dots, p_1, p_2, \dots) = \frac{N!}{k_1! \cdot k_2! \cdot \dots} p^{k_1} p^{k_2} \dots \quad (2.12)$$

Valore di aspettazione e varianza rimangono le medesime per ogni singolo bin, solo che i conteggi k_i non sono variabili aleatorie indipendenti tra loro, in questo caso si ha la **covarianza**:

$$cov[k_i, k_j] = -N \cdot p_i \cdot p_j \quad (2.13)$$

2.6 Distribuzione di Poisson

La distribuzione di Poisson è una distribuzione di probabilità discreta che esprime la probabilità che un numero di di eventi indipendenti si verifichi in un certo intervallo di spazio o tempo fissato con una frequenza media costante. il verificarsi dell'evento a un tempo t non influenza il verificarsi del medesimo in un intervallo di tempo successivo o precedente al tempo t . La probabilità di contare k eventi in un intervallo di tempo unitario è:

$$Poiss(k, \lambda) = \frac{e^{-\lambda} \cdot \lambda^k}{k!} \quad (2.14)$$

lo spazio campione rispetto alla quale è definita è tutto \mathbb{N}

2.6.1 Proprietà della distribuzione di Poisson

Il valore di aspettazione e la varianza della distribuzione di Poisson sono:

$$E[k] = \sum_{k=0}^{\infty} k \cdot \frac{e^{-\lambda} \cdot \lambda^k}{k!} = \lambda \quad (2.15)$$

$$V[k] = \sum_{k=0}^{\infty} (k - \lambda)^2 \cdot \frac{e^{-\lambda} \cdot \lambda^k}{k!} = \lambda \quad (2.16)$$

si nota che all'aumentare della media aumenta anche la varianza.

L'indice di simmetria e kurtosi sono dati da:

$$\gamma_1 = \frac{1}{\sqrt{\lambda}} \quad (2.17)$$

$$\gamma_2 = \frac{1}{\lambda} \quad (2.18)$$

per $\lambda \rightarrow \infty$ la distribuzione di Poisson diventa più simmetrica. Inoltre la distribuzione di Poisson gode della proprietà di **riproduttività**.

2.6.2 Da eventi Poissoniani all'esponenziale

Il modello su cui si basa la distribuzione di Poisson è dato dalle seguenti assunzioni:

- la probabilità che un evento avvenga in un intervallo di tempo $[t, t+dt]$ è data da $p \cdot dt$, dove p è un valore costante e indipendente da t ;
- la probabilità di avere più di un evento nell'intervallo dx è nulla.
- gli eventi sono indipendenti tra loro.

Definiamo $E = \{\text{nessun evento avvenuto in } [0, t+dt]\}$ e la funzione $q(t) = \{\text{probabilità di nessun evento in } [0, t]\}$, posta I come l'informazione a priori (ex: tempo di decadimento medio) si ha che la probabilità che si verifichi E essendosi già verificato I è:

$$P(E|I) \equiv q(t + dt) \quad (2.19)$$

Dove q ha come argomento la lunghezza degli intervalli di tempo; possiamo riformulare E come l'intersezione di due eventi $E_1 = \{0 \text{ eventi in } [0, t]\}$ ed $E_2 = \{0 \text{ eventi in } [t, t+dt]\}$, dunque 2.19 può essere riscritta come:

$$P(E|I) = P(E_1, E_2|I) = P(E_1|I)P(E_2|E_1, I)$$

poichè per ipotesi E_1 ed E_2 sono indipendenti dato che il verificarsi di un evento in un intervallo di tempo non influenza il verificarsi del medesimo evento in un intervallo successivo si ha che:

$$P(E|I) = P(E_1|I)P(E_2|I) \quad (2.20)$$

dove $P(E|I) = q(t + dt)$, $P(E_1|I) = q(t)$ e $P(E_2|I) = 1 - p dt$. Dunque:

$$q(t + dt) = q(t) \cdot (1 - p dt) \Rightarrow \frac{q(t + dt) - q(t)}{dt} = -q(t)p \quad (2.21)$$

questo è equivalente all'equazione differenziale al primo ordine:

$$\frac{dq}{q} = -p \cdot dt \iff q(t) = q(0)e^{-pt}$$

dove $q(0)$ è determinata normalizzando la distribuzione di probabilità. Per eventi che occorrono in modo Poissoniano la probabilità di non avere eventi in un intervallo $[0, x]$ è data da una esponenziale.

Per comodità poniamo $q(0)=1$. Definiamo $A = \{1 \text{ evento in } [t_1, t_1+dt_1]\}$ ed $B = \{0 \text{ eventi in } [0, t_1]\}$ definiamo:

$$P(A, E_1|I) = P(A|I) \cdot P(E_1|I) = e^{-pt}(p \cdot dt) \quad (2.22)$$

ipotizzando di costruire una catena di eventi in cui si realizza e non si realizza un evento per tempi ordinati si ha che:

$$P(A, E|I) = e^{-pt_1} \cdot p \cdot dt_1 \cdot e^{-p(t_2-t_1)} \cdot p \cdot dt_2 \cdots = e^{-pt} p^n dt_1 \cdots dt_n \quad (2.23)$$

Per definire una probabilità che si realizzino n eventi in un intervallo di tempo $[0, t]$, a prescindere dal tempo preciso in cui avvengono, dobbiamo sommare le probabilità definite in (2.23):

$$P(n|p, t, I) = e^{-pt} p^n \int_0^{t_2} dt_1 \cdot \int_0^{t_3} dt_2 \cdots \int_0^t dt_n \quad (2.24)$$

il prodotto d'integrali ha come soluzione $\frac{t^n}{n!}$ e dunque:

$$P(n|p, t, I) = \frac{(pt)^n}{n!} e^{-pt} = \frac{\lambda^n}{n!} e^{-\lambda} \quad (2.25)$$

2.7 Distribuzione Esponenziale

La distribuzione esponenziale è la distribuzione dei tempi che intercorrono tra due eventi Poissoniani. Infatti consideriamo k il numero di eventi in un processo Poissoniano con frequenza p , avvenuti in un intervallo di tempo $[0, t]$. Allora si ha:

$$P(k, p) = \frac{(pt)^k}{k!} e^{-pt}$$

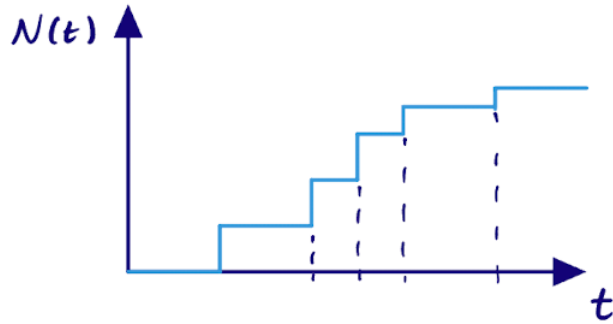
Sia $T > t$ il tempo in cui effettuiamo la prima misura, poichè la frequenza p è una costante indipendente dal tempo, se misuriamo la probabilità per 0 eventi in un tempo $t < T$ si avrà:

$$P(0|p, t) = \frac{(pt)^0}{0!} e^{-pt} = e^{-pt}$$

Se rileviamo il medesimo evento per $T > t$ avrà la stessa probabilità dell'evento avvenuto al tempo t :

$$P(0|p, T > t) = P(0|p, t) = e^{-pt}$$

Dunque la distribuzione di probabilità che intercorre da tra due eventi Poissoniani è esponenziale.



ed è definita da:

$$pdf(t, \lambda) = \lambda \cdot e^{-\lambda t} \quad (2.26)$$

dove $\tau = E[t]$ è il tempo medio tra due eventi consecutivi, e vale la relazione $\lambda = \frac{1}{\tau}$ dove λ è il parametro della Poissoniana che determina la frequenza media di eventi.

2.7.1 Proprietà della distribuzione esponenziale

Il valore di aspettazione e varianza sono date dalle quantità:

$$E[t] = \tau = \frac{1}{\lambda} \quad (2.27)$$

$$V[t] = \tau^2 = \frac{1}{\lambda^2} \quad (2.28)$$

mentre i parametri di simmetria della distribuzione dati da simmetria e kurtosi:

$$\gamma_1 = 2 \quad (2.29)$$

$$\gamma_2 = 6 \quad (2.30)$$

come si osserva dai momenti essendo γ_1 e γ_2 costanti la distribuzione esponenziale non può tendere ad una Gaussiana al crescere del numero di eventi.

2.7.2 Decadimento Radioattivo

Si consideri un campione di materiale radioattivo contenente N_0 nuclei al tempo $t=0$ per un singolo nucleo padre la probabilità di decadimento segue una distribuzione esponenziale il cui τ tempo medio di decadimento è dato dalla meccanica quantistica. Gli eventi di decadimento sono **indipendenti** tra di loro.

La variazione del numero dei nuclei in un intervallo di tempo Δt sarà proporzionale al numero di nuclei restanti:

$$-\frac{dN}{dt} = \lambda N \quad (2.31)$$

dove λ è la frequenza media di decadimento. Di conseguenza il decadimento radioattivo è definito dalla funzione:

$$N(t) = N_0 e^{-\lambda t} \quad (2.32)$$

che possiamo riscrivere come :

$$N(t) = N_0 e^{-\frac{t}{\tau}} \quad (2.33)$$

Se $\tau \gg t$ tempo di osservazione, la frequenza media dei decadimenti è costante nel tempo e di conseguenza la probabilità di osservare k decadimenti in un campione di N nuclei è data dalla distribuzione di Poisson con $\lambda = \frac{1}{\tau}$.

2.8 Comportamenti Asintotici delle distribuzioni

2.8.1 Proprietà asintotiche della Binomiale e Poissoniana

Binom($k|N, p$) \rightarrow Poisson($k, \mu = N \cdot p$)

La relazione tra distribuzione Binomiale e Poissoniana si costruisce assumendo di prendere un intervallo Δt e di dividerlo in N intervalli di dimensione dt tali per cui:

- in ciascun dt cada uno solo evento(tale ipotesi è verificata solo se il numeri di eventi è molto grande);
- la probabilità in un singolo intervallo dt è $p = \frac{\lambda}{N}$ ed è finita;
- gli eventi che cadono in ogni intervallo dt sono indipendenti tra loro.

Essendo gli eventi indipendenti tra loro per ipotesi ciascuno di essi segue una distribuzione di Bernoulli. Tali condizioni consentono di usare una distribuzione binomiale per descrivere eventi Poissoniani.

Per induzione si ha che per $k=0$ posto $\lambda = N \cdot p$

$$B\left(0|N, \frac{\lambda}{N}\right) = \left(1 - \frac{\lambda}{N}\right)^N \underset{N \rightarrow \infty}{=} \frac{\lambda^0}{0!} e^{-\lambda}$$

dunque la relazione per $k-1$ sarà verificata e:

$$B\left(k-1|N, \frac{\lambda}{N}\right) \underset{N \rightarrow \infty}{=} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda}$$

Allora si dimostra che dall'espressione binomiale:

$$\frac{B\left(k|N, \frac{\lambda}{N}\right)}{B\left(k-1|N, \frac{\lambda}{N}\right)} = \frac{Np - (k-1)p}{k(1-p)} \approx_{p \rightarrow 0} \frac{\lambda}{k}$$

dunque:

$$B\left(k|N, \frac{\lambda}{N}\right) = \frac{\lambda}{k} \cdot B\left(k-1|N, \frac{\lambda}{N}\right) \rightarrow \frac{\lambda^k}{k!} e^{-\lambda}$$

Si dimostra inoltre che data una distribuzione Binomiale $B(k|N, p) \rightarrow G(k, \mu, \sqrt{\mu})$ al crescere del numero di misure $N \rightarrow \infty$ per $\mu = N \cdot p$.

Vale anche il fatto che per una Poissoniana si ha $Poiss(k, \lambda) \rightarrow G(k, \lambda, \sqrt{\lambda})$ per $\lambda \rightarrow \infty$.

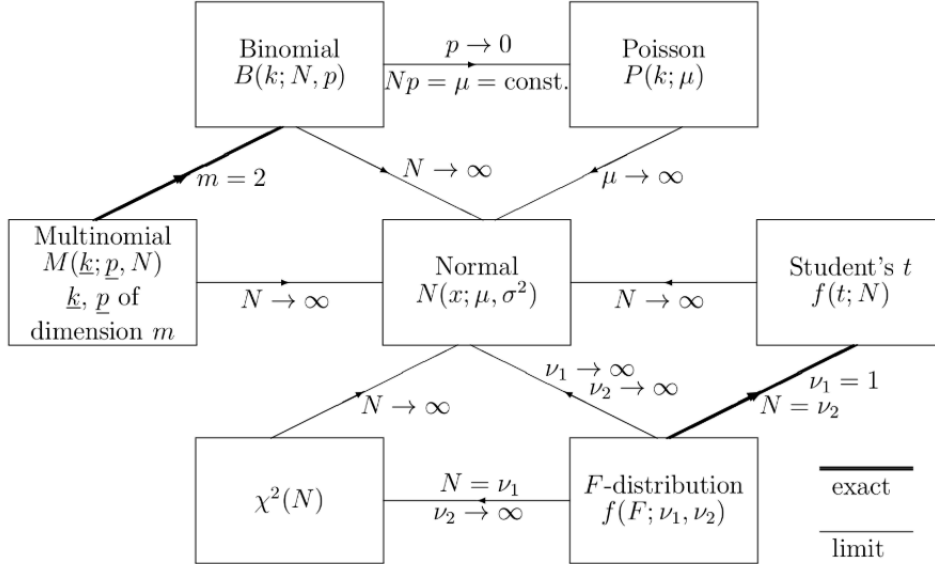


Figure 2.1: Mappa delle distribuzioni asintotiche

2.9 Distribuzione di χ^2

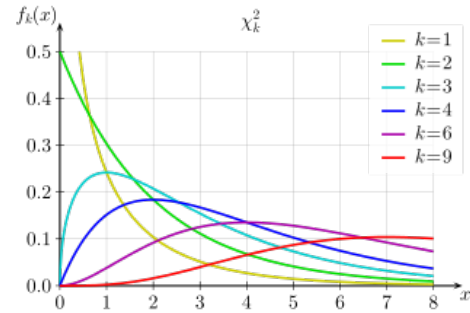
La distribuzione di χ^2 nasce nel contesto della modellazione di dati per la stima di parametri usando i metodi di maximum likelihood e minimi quadrati. Dato un campione di n punti $\{(x_i, y_i)\}_i^N$ e un modello $f(x, \underline{\theta})$, dipendente da m parametri $\underline{\theta}$ si valuta la bontà del modello nell'approssimare i dati usando la funzione di Q^2 definita come:

$$Q^2 = \sum_{i=1}^n \frac{[y_i - f(x_i, \underline{\theta})]^2}{\sigma_i^2} \quad (2.34)$$

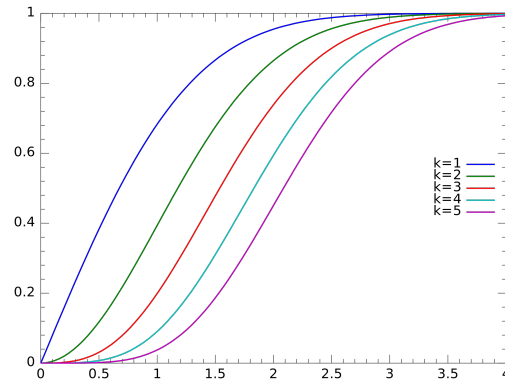
dove le σ_i^2 sono gli errori sulle n misure di y_i .

Per le misure y che seguono una pdf Gaussiana, la pdf del χ^2 è data distribuzione:

$$\chi^2 = \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} z^{(\frac{n}{2}-1)} e^{-\frac{z}{2}}$$



dove n prende il nome di **numero di gradi di libertà** e $\Gamma(x)$ è la funzione gamma.



La distribuzione di χ^2 possiede una distribuzione cumulativa e gode delle seguenti proprietà:

- **media:** $E[n] = n$
- **Varianza:** $V[n] = 2n$
- è riproduttiva
- **moda:** $n-2$

χ^2 ridotto

Si definisce χ^2 ridotto data dal rapporto $\frac{\chi^2}{n}$ e la sua media è pari a 1.

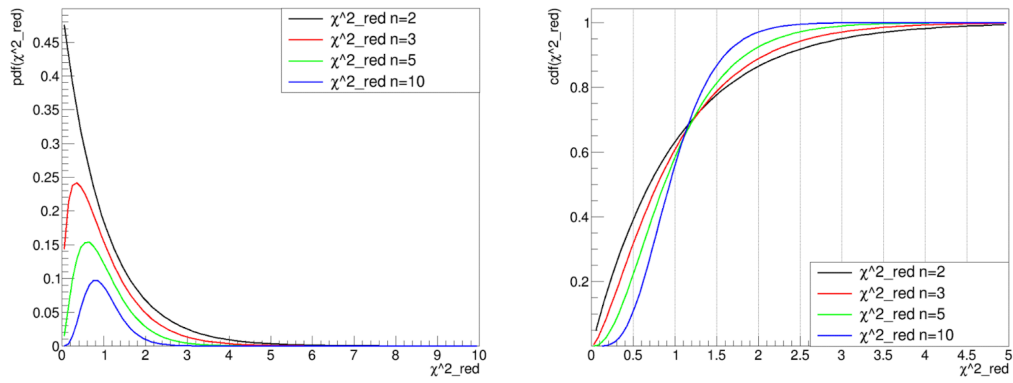


Figure 2.2: Distribuzione di χ^2 ridotto e rispetti distribuzione cumulativa.