Distribuzioni di probabililtà

Bernulli

$$X \sim B(p)$$

La distribuzione di Bernulli è una distribuzione di probabilità discreta che descrive un esperimento con due possibili risultati, successo o fallimento, con probabilità di successo p e probabilità di fallimento $q = 1 - p$.

	Valore atteso	Varianza
Bernulli	$\mathbb{E}(X) = p$	$Var(X) = p(1-p)$

Binomiale

$$X \sim B(n, p)$$

La distribuzione binomiale è una distribuzione di probabilità discreta che descrive il numero di successi in una sequenza di n esperimenti indipendenti, ciascuno con probabilità di successo p.

Per calcolare la probabilità di ottenere k successi in n esperimenti possiamo usare la formula:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Il coefficiente binomiale $\binom{n}{k}$ è definito come:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

	Valore atteso	Varianza
Binomiale	$\mathbb{E}(X) = np$	$Var(X) = np(1-p)$

Geometrica (successo dopo k tentativi)

$$X \sim G(p)$$

La distribuzione geometrica è una distribuzione di probabilità discreta che descrive il numero di tentativi necessari per ottenere il primo successo in una sequenza di esperimenti indipendenti, ciascuno con probabilità di successo p.

La probabilità di ottenere il primo successo al tentativo k è data da:

$$P(X = k) = (1-p)^{k-1} p$$

Questa descrive la probabilità di fallire $k - 1$ volte e poi avere successo al tentativo k, quindi avere 0 successi fino a $k - 1$.

NB: questa variabile **parte da 1**, quindi $k \geq 1$.

	Valore atteso	Varianza
Geometrica	$\mathbb{E}(X) = \frac{1}{p}$	$Var(X) = \frac{1-p}{p^2}$

Poisson - evento raro

$$X \sim P(\lambda)$$

In questo caso usiamo la Poisson per approssimare la binomiale quando si hanno tanti tentativi, quindi con un fattoriale difficile da computare.

NB: si usa quando $n > 50, np < 5$

La probabilità di ottenere k successi si calcola con:

$$P(X=k)=\frac{e^{-\lambda}\lambda^k}{k!} \quad \text{con } \lambda = np$$

	Valore atteso	Varianza
Poisson	$\mathbb{E}(X)=\lambda$	$Var(X)=\lambda$

Poisson - successi in tempi continui

$$X \sim P(\lambda t)$$

La distribuzione di Poisson può essere usata per modellare il numero di eventi in un intervallo di tempo t.

La probabilità di ottenere k successi in un intervallo di tempo t è data da:

$$P(X=k)=\frac{e^{-\lambda t}(\lambda t)^k}{k!}$$

Dove λ è il *numero di eventi su unità di tempo media*. Ad esempio 2 chiamate ogni 3 minuti, $\lambda=\frac{2}{3}$.

	Valore atteso	Varianza
Poisson	$\mathbb{E}(X)=\lambda t$	$Var(X)=\lambda t$

Poisson primo successo

$$X \sim P(\lambda)$$

La distribuzione di Poisson può essere usata per modellare il tempo necessario per ottenere il primo successo in un intervallo di tempo t.

La probabilità di ottenere il primo successo al tempo t è data da:

$$P(X=t)=e^{-\lambda t}$$

Questa descrive la probabilità di fallire fino al tempo $t-1$ e poi avere successo al tempo t.

	Valore atteso	Varianza
Poisson	$\mathbb{E}(X)=\frac{1}{\lambda}$	$Var(X)=\frac{1}{\lambda^2}$

Variabile esponenziale

$$X \sim Exp(\lambda)$$

La distribuzione esponenziale è una distribuzione di probabilità continua che descrive il tempo tra gli eventi in un processo di Poisson.

La funzione di probabilità di una variabile esponenziale è data da:

$$f(x)=\lambda e^{-\lambda x} \quad \text{con } x \geq 0$$

essendo che descrive il tempo non può essere minore di 0

La **Funzione di sopravvivenza** che descrive la probabilità che un oggetto si guasti in un intervallo $[0,t]$ è data da:

$$P(X \leq x)=1-e^{-\lambda x}$$

La probabilità che si **guasti dopo un tempo** t è data da:

$$P(X>t)=e^{-\lambda t}=S(x)$$

Proprietà assenza di memoria

La distribuzione esponenziale gode della proprietà di assenza di memoria, cioè la probabilità che un evento si verifichi nei prossimi t minuti è indipendente dal tempo trascorso finora.

$$P(X > x_0 + x | X > x_0) = P(X > x)$$

Questa equazione ci dice quindi che anche se sappiamo che è passato un tempo x_o non condiziona la probabilità, ma rimane sempre la stessa.

Esempio:
Quale è la probabilità che la riparazione richieda almeno 3 ore sapendo che ne richiede più di 2?

$$P(X > 3 | X > 2) = P(X > 1) = e^{-\lambda}$$

Quindi la probabilità è la stessa.
Importante: quindi quando calcoliamo la probabilità $x = |x - x_0|$.

Legge dei grandi numeri

La legge dei grandi numeri afferma che la media di un numero grande di campioni di una variabile aleatoria indipendente e identicamente distribuita converge al valore atteso della variabile aleatoria.

Teorema del limite centrale

Il teorema del limite centrale afferma che la somma di un numero grande di variabili aleatorie indipendenti e identicamente distribuite converge a una variabile aleatoria normale.

Media campionaria

La media campionaria è la media di un campione di n osservazioni di una variabile aleatoria. La media campionaria è una variabile aleatoria e il suo valore atteso è uguale al valore atteso della variabile aleatoria originale.

	Valore atteso	**Varianza**
Media campionaria	$\mathbb{E}(\bar{X}) = \mu$	$Var(\bar{X}) = \frac{\sigma^2}{n}$

Distribuzione normale standard

$$X \sim N(\mu, \sigma^2)$$

Nel caso di una normale standard, la media è 0 e la varianza è 1. Quindi la scriviamo come: $X \sim N(0, 1)$.

La distribuzione normale è una distribuzione di probabilità continua che è simmetrica rispetto alla media μ e ha una deviazione standard σ.

La il calcolo della probabilità si indica con il simbolo greco Φ:

$$P(X \leq b) = \Phi(b)$$

	Valore atteso	**Varianza**
Normale	$\mathbb{E}(X) = \mu$	$Var(X) = \sigma^2$

Standardizzazione

Per calcolare la probabilità di una normale non standardizzata dobbiamo standardizzarla, quindi con media 0 e varianza 1. Per farlo usiamo la formula:

$$Z = \frac{X - \mu}{\sigma}$$

Quindi possiamo calcolare la probabilità di una normale non standardizzata come:

$$P(a \leq X \leq b) = P\left(\frac{a - \mu}{\sigma} \leq Z \leq \frac{b - \mu}{\sigma}\right)$$

Teorema del limite centrale per una Bernulli

Il teorema del limite centrale afferma che la somma di un numero grande di variabili aleatorie indipendenti e identicamente distribuite converge a una variabile aleatoria normale.
Questo lo possiamo applicare anche alla distribuzione di una Bernulli per facilitare i conti.

NB: per poter usare l'approssimazione deve valere $np > 5$ e $n(1-p) > 5$.

Approssimando avremo:

$$X \sim B(n,p) \approx N(np, np(1-p))$$

Invece la standaridizzazione diventa:

$$Z = \frac{X - np}{\sqrt{np(1-p)}}$$

Essendo che passiamo da una distribuzione discreta a una continua, dobbiamo fare una correzione di continuità, quindi usiamo le seguenti regole con X **prima di standarizzare**:

$$\begin{cases} P(X > \phi) & \text{diventa} \quad P(X > x + 0.5) \\ P(X \leq x) & \text{diventa} \quad P(X \leq x + 0.5) \\ \text{Altrimenti}(\geq, <) & \text{diventa} \quad \quad P(x - 0.5) \end{cases}$$

Formulario distribuzioni

		$\mathbb{E}$	Var	**Funzione di probabilità**		**Note**
Bernulli	$B(p)$	p	$p(1-p)$	$p^k(1-p)^{1-k}$		-
Binomiale	$B(n,p)$	np	$np(1-p)$	$\binom{n}{k}p^k(1-p)^{n-k}$		-
Geometrica	$G(p)$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$(1-p)^{k-1}p$		parte da 1
Poisson - Binom	$P(\lambda)$	λ	λ	$\frac{e^{-\lambda}\lambda^k}{k!}$	$\lambda = np$	$n > 50, np < 5$
Poisson - temp	$P(\lambda t)$	λ	λ	$\frac{e^{-\lambda}\lambda^k}{k!}$	λ =successi su unità di tempo, t: tempo esercizio (prob 2 chiamate nei prox 4min(t=4))	es 3 chiam ogni 5 min:$\lambda = \frac{3}{5}$
Esponenziale	$Exp(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$P(X \leq x) = 1 - \lambda e^{-\lambda x}$	$S(X > x) = e^{-\lambda x}$	Soppravv - guasto dopo t
Normale	$N(\mu, \sigma^2)$	μ	σ^2	$Z = \frac{X-\mu}{\sigma}$		Normalizzazione
Normale per Bernulli	$N(np, np(1-p))$	μ	σ^2	$Z = \frac{X-np}{\sqrt{np(1-p)}}$		$>, \leq \rightarrow x + 0.5$ $\geq, < \rightarrow x - 0.5$ $esattamente \rightarrow x \pm 0.5$
---	---	---	---	---	---	---
Varianza	$S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$	Deviazione standard: $S = \sqrt{S^2}$				

Nota:
Deviazione standard $\rightarrow \sigma$
Varianza $\rightarrow \sigma^2$