Capitolo Nr.1

Appunti di probabilità

1.1 Combinatoria

	Semplici	**Con ripetizioni**
Permutazioni	$n!$	$\frac{n!}{m!}$
Disposizioni	$\frac{n!}{(n-m)!}$	n^k
Combinazioni	$\frac{n!}{m!\cdot(n-m)!} = \binom{n}{m}$	$\frac{(n+m-1)!}{m!\cdot(n-1)!}$

- *Permutazioni semplici*: in quanti modi posso riordinare n oggetti?
- *Permutazioni con ripetizioni*: in quanti modi posso riordinare n oggetti se un oggetto si ripete m volte?
- *Disposizioni semplici*: in quanti modi posso ordinare m oggetti scelti tra gli n di un insieme?
- *Disposizioni con ripetizioni*: in quanti modi posso disporre n oggetti in k cassetti?
- *Combinazioni semplici*: in quanti modi posso scegliere, in modo non ordinato, m oggetti tra gli n di un insieme?
- *Combinazioni con ripetizioni*: in quanti modi posso scegliere, in modo non ordinato, m oggetti tra gli n di un insieme, se gli oggetti da prendere sono più di quelli tra cui scegliere?

1.2 Probabilità condizionata e indipendenza

Sia $(\Omega, \mathcal{F}, P)$ uno spazio di probabilità e siano $F, E \in \mathcal{F}$ con $P(F) \neq 0$ eventi su quello spazio. La probabilità di E condizionato F è:

$$P(E|F) := \frac{P(E \cap F)}{P(F)}$$

Teorema di fattorizzazione o delle probabilità totali Dato uno spazio di probabilità $(\Omega, \mathcal{F}, P)$ e una sua partizione $\{E_i\}_{i \in I}$ tale che $\bigcup_{i \in I} E_i = \Omega$ e $P(E_i) \neq 0$, allora vale:

$$\forall E \in \mathcal{F} \ P(E) = \sum_{i \in I} P(E \cap E_i) = \sum_{i \in I} P(E|E_i) \cdot P(E_i)$$

Teorema di Bayes

(i) Se $P(E) \neq 0 \neq P(F)$ vale:

$$P(E|F) = \frac{P(F|E) \cdot P(E)}{P(F)}$$

(ii) Se ho una partizione $\{E_i\}_{i \in I}$ tale che $\bigcup_{i \in I} E_i = \Omega$ e $P(E_i) \neq 0$, allora vale:

$$P(E_j|F) = \frac{P(F|E_j) \cdot P(E_j)}{\sum_{i \in I}(P(F|E_i) \cdot P(E_i))}$$

Indipendenza Due eventi $F, E \in \mathcal{F}$ sono indipendenti se e sole se:

$$P(E \cap F) = P(E) \cdot P(F)$$

Indipendenza condizionata Dati due eventi $E_1, E_2 \in \mathcal{F}$ in uno spazio di probabilità $(\Omega, \mathcal{F}, P)$ e fissato un terzo evento $F \in \mathcal{F}$, E_1 ed E_2 sono indipendenti condizionatamente a F se:

$$P(E_1 \cap E_2|F) = P(E_1|F) \cdot P(E_2|F)$$

1.3 Variabili aleatorie

1.3.1 Variabili aleatorie discrete

- *Bernoulliane* $[X \sim bin(1, p)]$: è una *variabile aleatoria indicatrice* di un evento $E =$ "*successo*" che si verifica con probabilità p.

$$\varphi_X(x) = \begin{cases} p & \text{se } x = 1 \\ 1-p & \text{se } x = 0 \\ 0 & \text{altrimenti} \end{cases} \text{ e } F_X(x) = \begin{cases} 0 & \text{se } x < 0 \\ 1-p & \text{se } 0 \leq x < 1 \\ 1 & \text{se } x \geq 1 \end{cases}$$

- *Binomiali* $[X \sim bin(n, p)]$: conta il numero di successi che si verificano in n tentativi.

  $\varphi_X(k)$: indica con che probabilità si verificano k successi

$$\varphi_X(k) = \begin{cases} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} & \text{se } k \in \{0, .., n\} \\ 0 & \text{altrimenti} \end{cases}$$

$$F_X(x) = \sum_{k=0}^{\lfloor x \rfloor} \varphi_X(k) = \sum_{k=0}^{min\{\lfloor x \rfloor, n\}} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

- *Geometriche* $[X \sim geom(p)]$: indica l'istante precedente al primo successo.

  $\varphi_X(k)$: indica con che probabilità si verificano k insuccessi prima del primo successo

$$\varphi_X(k) = P(X = k) = \begin{cases} 0 & \text{se } k \notin \mathbb{N} \\ (1-p)^k \cdot p & \text{se } k \in \mathbb{N} \end{cases} \text{ e } F_X(x) = \begin{cases} 0 & \text{se } x < 0 \\ 1 - (1-p)^{\lfloor x \rfloor + 1} & \text{se } x \geq 0 \end{cases}$$

- *Binomiali negative* $[X \sim NB(n, p)]$: conta il numero di insuccessi precedenti all'n-esimo successo.

  $\varphi_X(k)$: indica con che probabilità si verificano k insuccessi prima dell'n-esimo successo

$$\varphi_X(x) = p^n \cdot (1-p)^k \cdot \binom{k+n-1}{k}$$

- *Ipergeometriche* [$X \sim hyp(k, m, n)$]: conta il numeri di oggetti di tipo a estratti, senza reimmissione, da un insieme di m oggetti di tipo a ed n di tipo b.

  $\varphi_X(k)$: indica con che probabilità vengono estratti k oggetti di tipo a tra gli m di tipo a e gli n di tipo b

$$\varphi_X(b) = \begin{cases} \frac{\binom{n}{b} \cdot \binom{n}{k-b}}{\binom{m+n}{k}} & \text{se } max\{0, k-n\} \leq b \leq min\{k, m\} \\ 0 & \text{altrimenti} \end{cases}$$

- *Poissoniane* [$X \sim pois(\lambda)$]: si usa quando si ha una successione di *ipergeometriche* di parametri a_i e b_i che tende a una *binomiale* con probabilità di successo α.

  $\varphi_X(k)$: se, ad esempio, si stanno considerando i morti giornalieri sul lavoro, e in media, ne muoiono $3(=\lambda)$ ogni giorno, indica la probabilità che un giorno ne muoiano k

$$\varphi_X(k) = \begin{cases} \frac{\lambda^k}{k!} \cdot e^{-\lambda} & \text{se } k \in \mathbb{N} \\ 0 & \text{altrimenti} \end{cases}$$

1.3.2 Variabili aleatorie assolutamente continue

- *Uniformi* [$X \sim unif(a, b)$]: hanno densità costante in $[a, b]$.

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{se } x \in (a, b) \\ 0 & \text{altrimenti} \end{cases} \text{ e } F_X(x) = \begin{cases} 0 & \text{se } x \leq a \\ \frac{x-a}{b-a} & \text{se } a < x \leq b \\ 1 & \text{se } x > b \end{cases}$$

- *Esponenziali* [$X \sim exp(\lambda)$]: sono l'equivalente continuo delle geometriche.

$$f_X(x) = \begin{cases} 0 & \text{se } x < 0 \\ c \cdot e^{-\lambda x} & \text{se } x \geq 0 \end{cases} \text{ e } F_X(x) = \begin{cases} 0 & \text{se } x < 0 \\ 1 - e^{-\lambda x} & \text{se } x \geq 0 \end{cases}$$

- *Gaussiane* [$X \sim \mathcal{N}(\mu, \sigma)$]: i parametri sono rispettivamente media e deviazione standard.

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \text{ e } F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

  Se $X \sim \mathcal{N}(0, 1)$, X è una normale standard e la funzione di ripartizione è:

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

- *Chi quadro* [$X \sim \chi^2(n)$]: sono definite come somma di Gaussiane standard indipendenti. Il parametro n indica i gradi di libertà, ovvero il numero di Gaussiane che sono state sommate per realizzare la Chi quadro.

- *t di Student* [$X \sim t(n)$]: X è definita come

$$X = \frac{Z}{\sqrt{\frac{W}{n}}} \text{ con } Z \sim \mathcal{N}(0, 1) \text{ e } W \sim \chi^2(n)$$

1.3.3 Riproducibilità

Prop. La somma di n *Bernoulliane* indipendenti e identicamente distribuite di parametro p è distribuita come una *binomiale* di parametri n e p. In particolare, se $X_i \sim bin(1,p)$ $\sum_{i=1}^{n} X_i \sim bin(n,p)$.

Prop. La famiglia delle *binomiali* a parametro p fissato è riproducibile. In particolare, se $X \sim bin(n,p)$ e $Y \sim bin(m,p)$ sono indipendenti $X+Y \sim bin(n+m,p)$.

Prop. La famiglia delle *binomiali negative* a parametro p fissato è riproducibile. In particolare, se $X \sim NB(n,p)$ e $Y \sim NB(m,p)$ sono indipendenti allora $X+Y \sim NB(n+m,p)$.

Prop. La somma di n *geometriche* indipendenti e identicamente distribuite di parametro p è distribuita come una *binomiale negativa* di parametri n e p. In particolare, se $X_i \sim geom(p)$ $\sum_{i=1}^{n} X_i \sim NB(n,p)$.

Prop. La famiglia delle *Poissoniane* è riproducibile. In particolare, se $X \sim pois(\lambda_1)$ e $Y \sim pois(\lambda_2)$ sono indipendenti allora $X+Y \sim pois(\lambda_1+\lambda_2)$.

Prop. Sia $Z \sim \mathcal{N}(0,1)$. X è una *normale* o *Gaussiana* di parametri $\mu, \sigma \in \mathbb{R}$ con $\sigma > 0$ se:

$$X = \sigma * Z + \mu$$

In particolare, $X \sim \mathcal{N}(\mu, \sigma)$.

Prop. La famiglia delle *Chi quadro* è riproducibile. In particolare, se $X \sim \chi^2(n)$ e $Y \sim \chi^2(m)$ sono indipendenti allora $X+Y \sim \chi^2(n+m)$.

1.4 Indicatori per variabili aleatorie

1.4.1 Valore atteso

Valore atteso di variabili aleatorie

- X *v.a. discreta*: $E[X] = \sum_{x \in \mathcal{R}_X} x \cdot \varphi_X(x)$
- X *v.a. assolutamente continua*: $E[X] = \int_{-\infty}^{+\infty} x \cdot f_X(x)dx$

Valore atteso di trasformazioni di variabili aleatorie Sia $Y = g(X)$, vale:

- X *v.a. discreta*: $E[Y] = \sum_{x \in \mathcal{R}_X} g(x) \cdot \varphi_X(x)$
- X *v.a. assolutamente continua*: $E[Y] = \int_{-\infty}^{+\infty} g(x) \cdot f_X(x)dx$

Valore atteso di 2-vettori Sia (X,Y) un *2-vettore* e sia $Z = g(X,Y)$, vale:

- *2-vettore discreto*: $E[Z] = \sum_{x \in \mathcal{R}_X} \sum_{y \in \mathcal{R}_Y} g(x,y) \cdot \varphi_{X,Y}(x,y)$
- *2-vettore assolutamente continuo*: $E[Z] = \int \int_{\mathbb{R}^2} g(x,y) \cdot f_{X,Y}(x,y)dx\, dy$
- *2-vettore misto*:
  - X *discreta*, Y *assolutamente continua*: $E[Z] = \sum_{x \in \mathcal{R}_X} \int_{\mathbb{R}} g(x,y) \cdot f_{X,Y}(x,y)dy$
  - X *assolutamente continua*, Y *discreta*: $E[Z] = \int_{\mathbb{R}} \sum_{y \in \mathcal{R}_Y} g(x,y) \cdot f_{X,Y}(x,y)dx$

Proprietà del valore atteso

(i) *Linearità*: $E[a \cdot X + b] = a \cdot E[X] + b$

(ii) *Prodotto di v.a. indipendenti*: $E[X \cdot Y] = E[X] \cdot E[Y]$

(iii) *Monotonia*:

- $X = 0 \Rightarrow E[X] = 0$
- $X \geq 0 \Rightarrow E[X] \geq 0$

1.4.2 Momenti, varianza, deviazione standard

Momento n-esimo Siano $n \in \mathbb{N}$ con $n > 0$ e X variabile aleatoria:

- *Momento n-esimo*: $E[X^n]$
- *Momento n-esimo centrato*: $E[(X - E[X])^n]$

Varianza di variabili aleatorie Sia X una variabile aleatoria. La varianza di X è:

$$Var[X] = E[X^2] - (E[X])^2$$

Proprietà della varianza

(i) $Var[X] \geq 0$; $Var[X] = 0 \Leftrightarrow X = c \; \forall c \in \mathbb{R}$

(ii) $Var[a \cdot X + b] = a^2 \cdot Var[X]$

Deviazione standard Sia X una variabile aleatoria. La deviazione standard di X è:

$$\sigma(X) = \sqrt{Var[X]}$$

1.4.3 Covarianza e correlazione

Covarianza Siano X e Y due variabili aleatorie. La covarianza di X e Y è:

$$Cov[X, Y] = E[(X - E[X]) \cdot (Y - E[Y])] = E[X, Y] - E[X] \cdot E[Y]$$

Se $X = Y$, vale:

$$Cov[X, X] = E[(X - E[X])^2] = Var[X]$$

Proprietà della covarianza

(i) Se X e Y sono indipendenti allora $Cov[X, Y] = 0$

(ii) $Var[X + Y] = Var[X] + Var[Y] + 2 * Cov[X, Y]$

(iii) *Simmetria*: $Cov[X, Y] = Cov[Y, X]$

(iv) *Linearità*: $Cov[a * x + b * Y, Z] = a * Cov[X, Z] + b * Cov[Y, Z]$

(v) *Bilinearità*: Siano $(a_i)_{i=1}^n$ e $(b_j)_{j=1}^m$ *vettori reali* e $(X_i)_{i=1}^n$ e $(Y_j)_{j=1}^m$ *vettori aleatori*, vale:

$$Cov\left[\sum_{i=1}^{n} a_i * X_i, \sum_{j=1}^{m} b_j * Y_j\right] = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i * b_j * Cov[X_i, Y_j]$$

Correlazione e scorrelazione Due variabili aleatorie X e Y sono scorrelate se e solo se $Cov[X, Y] = 0$, altrimenti sono correlate.

Coefficiente di correlazione Date due *variabili aleatorie* X e Y si dice *coefficiente di correlazione* il numero:

$$\rho(X, Y) = corr[X, Y] = \frac{Cov[X, Y]}{\sqrt{Var[X] * Var[Y]}}$$

1.4.4 Moda, mediana, mediana impropria

Moda Chiamiamo *moda* di una *variabile aleatoria* X il un numero $x \in \mathcal{R}_X$ tale che:

- Se X è *discreta* φ_X è massima in x, cioè $x \in argmax_y \varphi_X(y)$
- Se X è *assolutamente continua* f_X è massima in x, cioè $x \in argmax_y f_X(y)$

Intuitivamente, la moda di una variabile aleatoria è il valore più probabile.

Mediana Si dice *mediana* di una *variabile aleatoria* X un numero m_X tale che:

$$P(X \leq m_X) = P(X \geq m_X)$$

Mediana impropria Si dice *mediana impropria* un numero reale $\tilde{m}_X$ tale che:

$$P(X \leq \tilde{m}_X) \geq \frac{1}{2} \wedge P(X \geq \tilde{m}_X \geq \frac{1}{2})$$

1.4.5 Indicatori per modelli aleatori

- *Bernoulliane* $[X \sim bin(1, p)]$: $E[X] = p$, $Var[X] = p \cdot (1 - p)$
- *Binomiali* $[X \sim bin(n, p)]$: $E[X] = n \cdot p$, $Var[X] = n \cdot p \cdot (1 - p)$
- *Geometriche* $[X \sim geom(p)]$: $E[X] = \frac{1-p}{p}$, $Var[X] = \frac{1-p}{p^2}$
- *Binomiali negative* $[X \sim NB(n, p)]$: $E[X] = n \cdot \frac{1-p}{p}$, $Var[X] = n \cdot \frac{1-p}{p^2}$
- *Ipergeometriche* $[X \sim hyp(k, m, n)]$: $E[X] = k \cdot \frac{m}{m+n}$
- *Poissoniane* $[X \sim pois(\lambda)]$: $E[X] = \lambda$, $Var[X] = \lambda$
- *Uniformi* $[X \sim unif(a, b)]$: $E[X] = \frac{b+a}{2}$, $Var[X] = \frac{(b-a)^2}{12}$
- *Esponenziali* $[X \sim exp(\lambda)]$: $E[X] = \frac{1}{\lambda}$, $Var[X] = \frac{1}{\lambda^2}$
- *Gaussiane* $[X \sim \mathcal{N}(\mu, \sigma)]$: $E[X] = \mu$, $Var[X] = \sigma^2$
- *Chi quadro* $[X \sim \chi^2(n)]$: $E[X] = n$, $Var[X] = 2n$
- *t di Student* $[X \sim t(n)]$: $E[X] = 1$, $Var[X] = \frac{2}{n}$

1.4.6 Disuguaglianze

Markov sia X una *variabile aleatoria* non negativa. Per ogni $a > 0$ vale:

$$P(X \geq a) \leq \frac{E[X]}{a}$$

Chebychev Sia X una *variabile aleatoria*. Per ogni $a > 0$ vale:

$$P(|X - E[X]| \geq a) \leq \frac{Var[X]}{a^2}$$

Capitolo Nr.2

Appunti di statistica

2.1 Stimatori

Correttezza e bias Uno stimatore Θ di ϑ è:

- *Corretto* se $E[\Theta] = \vartheta$
- *Scorretto* se $E[\Theta] \neq \vartheta$ e il valore $E[\Theta] - \vartheta$ è detto bias

Errore quadratico medio L'errore quadratico medio di uno stimatore Θ per ϑ è il valore:

$$MSE[\Theta] = E[(\Theta - \vartheta)^2] = Var[\Theta] + (bias)^2$$

Consistenza Uno stimatore Θ di ϑ è consistente se Θ_n converge in probabilità a ϑ, ovvero se:

$$\Theta_n \xrightarrow[n\to+\infty]{P} \vartheta$$

Consistenza in media quadratica Uno stimatore Θ di ϑ è consistente in media quadratica se Θ_n converge in media quadratica a ϑ, ovvero se:

$$\Theta_n \xrightarrow[n\to+\infty]{L^2} \vartheta$$

NB. La consistenza in media quadratica implica:

$$\lim_{n\to+\infty} E[(\Theta_n - \vartheta)^2] = 0$$

NB. Poiché la convergenza in media quadrati implica convergenza in probabilità, la consistenza in media quadratica implica consistenza, quindi uno stimatore è consistente se e solo se:

$$MSE[\Theta_n] = Var[\Theta_n]$$

Alcuni stimatori Sia $(X_1, \ldots, X_n)$ un campione di variabili aleatorie indipendenti e identicamente distribuite. I seguenti sono stimatori corretti e consistenti per quel campione:

- *Media campionaria*: è uno stimatore della media μ

$$\hat{\mu} = \bar{X} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$$

- *Varianza campionaria a media nota*: è uno stimatore della varianza σ a media μ nota

$$S_*^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} (X_i - \mu)^2$$

- *Varianza campionaria a media ignota*: è uno stimatore della varianza σ a media μ ignota

$$S^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \hat{\mu})^2$$

2.2 Funzioni ancillari

Funzione ancillare Chiamiamo funzione ancillare per un parametro ϑ una variabile aleatoria la cui legge sia nota a priori e che dipenda dai dati, da parametri noti e da ϑ, unico parametro non noto.

Alcune funzioni ancillari per Gaussiane

- Funzione ancillare per il parametro μ a varianza σ^2 nota:

$$\frac{\hat{\mu} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0, 1)$$

- Funzione ancillare per il parametro μ a varianza σ^2 ignota:

$$\frac{\hat{\mu} - \mu}{\sqrt{\frac{S^2}{n}}} \sim t(n-1)$$

- Funzione ancillare per la varianza σ^2 a speranza μ nota:

$$\frac{S_*^2}{\sigma^2} \cdot n = \sum_{i=1}^{n} (\frac{X_i - \mu}{\sigma})^2 \sim \chi^2(n)$$

- Funzione ancillare per la varianza σ^2 a speranza μ ignota:

$$\frac{S^2}{\sigma^2} \cdot (n-1) \sim \chi^2(n-1)$$

2.3 Costruire stimatori

2.3.1 Stimatori col metodo dei momenti

Sia $X = (X_1, \ldots, X_n)$ un campione di dimensione n di variabili aleatorie indipendenti e identicamente distribuite. In base alla distribuzione di un generico X_i, vale:

- $X_i \sim \mathcal{N}(\mu, \sigma)$: $\hat{\vartheta}_{mom} = (\hat{\mu}_{mom}, \hat{\sigma}^2_{mom})$, σ^2 ignota

$$\begin{cases} \hat{\mu}_{mom} = \hat{\mu} = \bar{X} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i \\ \hat{\sigma}^2_{mom} = \frac{n-1}{n} \cdot S^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - \mu)^2 \end{cases}$$

2.3.2 Stimatori di massima verosimiglianza

Sia $X = (X_1, \dots, X_n)$ un campione di dimensione n di variabili aleatorie indipendenti e identicamente distribuite. In base alla distribuzione di un generico X_i, vale:

- $X_i \sim bin(1, p)$: $\hat{\vartheta}_{MLE} = \hat{p}_{MLE}$

$$\hat{p}_{MLE} = \bar{X} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i$$

- $X_i \sim pois(\lambda)$: $\hat{\vartheta}_{MLE} = \hat{\lambda}_{MLE}$

$$\hat{\lambda}_{MLE} = \bar{X}$$

- $X_i \sim unif[-a, a]$: $\hat{\vartheta}_{MLE} = \hat{a}_{MLE}$

$$\hat{a}_{MLE} = \max\left(|\min_i(x_i|, |\max_i(x_i)|\right)$$

- $X_i \sim unif[a, b]$: $\hat{\vartheta}_{MLE} = (\hat{a}_{MLE}, \hat{b}_{MLE})$

$$\begin{cases} \hat{a}_{MLE} = \min_i(x_i) \\ \hat{b}_{MLE} = \max_i(x_i) \end{cases}$$

- $X_i \sim exp(\lambda)$: $\hat{\vartheta}_{MLE} = \hat{\lambda}_{MLE}$

$$\hat{\lambda}_{MLE} = \bar{X}^{-1}$$

- $X_i \sim \mathcal{N}(\mu, \sigma)$: $\hat{\vartheta}_{MLE} = (\hat{\mu}_{MLE}, \hat{\sigma}^2_{MLE})$, σ^2 ignota

$$\begin{cases} \hat{\mu}_{MLE} = \hat{\mu} = \bar{X} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i \\ \hat{\sigma}^2_{MLE} = \frac{n-1}{n} \cdot S^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - \mu)^2 \end{cases}$$

2.4 Intervalli di confidenza

2.4.1 Intervalli bilaterali

ϑ	**Note**	**Int. bilaterale**
μ	σ^2 nota	$\bar{X}_n \pm \Phi^{-1}(1 - \frac{\alpha}{2})\sqrt{\frac{\sigma^2}{n}}$
μ	σ^2 ignota	$\bar{X}_n \pm F^{-1}_{t_{n-1}}(1 - \frac{\alpha}{2})\sqrt{\frac{S^2}{n}}$
σ^2	μ nota	$\left(\frac{S^2_{*n} n}{F^{-1}_{\chi^2_n}(1-\frac{\alpha}{2})}, \frac{S^2_{*n} n}{F^{-1}_{\chi^2_n}(\frac{\alpha}{2})}\right)$
σ^2	μ ignota	$\left(\frac{S^2_n (n-1)}{F^{-1}_{\chi^2_{n-1}}\left(1-\frac{\alpha}{2},\right)}, \frac{S^2_n (n-1)}{F^{-1}_{\chi^2_{n-1}}\left(\frac{\alpha}{2},\right)}\right)$

2.4.2 Intervalli unilaterali

ϑ	**Note**	**Int. sinistro**	**Int. destro**
μ	σ^2 nota	$\left(-\infty, \bar{X}_n + \Phi^{-1}(1-\alpha)\sqrt{\frac{\sigma^2}{n}}\right)$	$\left(\bar{X}_n - \Phi_{-1}(1-\alpha)\sqrt{\frac{\sigma^2}{n}}, +\infty\right)$
μ	σ^2 ignota	$\left(-\infty, \bar{X}_n + F^{-1}_{t_{n-1}}(1-\alpha)\sqrt{\frac{S^2}{n}}\right)$	$\left(\bar{X}_n + F^{-1}_{t_{n-1}}(1-\alpha)\sqrt{\frac{S^2}{n}}, +\infty\right)$
σ^2	μ nota	$\left(0, \frac{S^2_{*n}n}{F^{-1}_{\chi^2_n}(\alpha)}\right)$	$\left(\frac{S^2_{*n}n}{F^{-1}_{\chi^2_n}(1-\alpha)}, +\infty\right)$
σ^2	μ ignota	$\left(0, \frac{S^2_n(n-1)}{F^{-1}_{\chi^2_{n-1}}(\alpha,)}\right)$	$\left(\frac{S^2_n(n-1)}{F^{-1}_{\chi^2_{n-1}}(1-\alpha,)}, +\infty\right)$

Larghezza di un intervallo Sia $X = (X_1, \ldots, X_n)$ un intervallo di variabili aleatorie indipendenti e identicamente distribuite come una Gaussiana di parametri μ e σ. La larghezza dell'intervallo di confidenza per la media μ al livello di confidenza $1-\alpha$ è:

$$l = 2\Phi^{-1}(1 - \frac{\alpha}{2})\sqrt{\frac{\sigma^2}{n}}$$

da cui è possibile ricavare la numerosità n del campione tale per cui, la larghezza di tale intervallo, non superi una larghezza l prefissata:

$$n = \left\lceil \frac{4(\Phi^{-1}(1-\frac{\alpha}{2}))^2\sigma^2}{l^2} \right\rceil$$

2.5 Intervalli di confidenza approssimati

2.5.1 Bernoulliane

Sia $X = (X_1, \ldots, X_n)$ un campione di variabili aleatorie Bernoulliane di parametro p. Vogliamo stimare p. Sia $Y_n = \sum_{i=1}^n X_i$ la variabile aleatoria che conta il numero di successi ($Y_n \sim bin(n,p)$). Poiché, nel caso delle Bernoulliane, p coincide con la media, $\hat{p} = \bar{X}_n = \frac{Y_n}{n}$ è uno stimatore della media e quindi di p. Vale, quindi:

$$P\left(\hat{p} - \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right) \approx 1-\alpha$$

Intervallo bilaterale

$$\hat{p} \pm \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Intervallo unilaterale sinistro

$$\left(0, \hat{p} + \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$$

Intervallo unilaterale destro

$$\left(\hat{p} - \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, 1\right)$$

Larghezza

$$l = 2\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \Rightarrow n = \left\lceil \frac{4(\Phi^{-1}(1-\frac{\alpha}{2}))^2}{l^2}\hat{p}(1-\hat{p}) \right\rceil \leq \left\lceil \frac{(\Phi^{-1}(1-\frac{\alpha}{2}))^2}{l^2} \right\rceil$$

Siccome, prima di iniziare a raccogliere i dati, non è possibile conoscere $\hat{p}$, si fanno delle misurazioni e poi si fa una prima stima di $\hat{p}$, che sarà quindi utilizzata per stimare la numerosità del campione.

2.5.2 Poissoniane

Sian $X = (X_1, \ldots, X_n)$ un campione di variabili aleatorie indipendenti e identicamente distribuite come Poissoniane di parametro λ. Vogliamo stimare λ. Poiché λ coincide con la media, $\bar{X}_n$ è uno stimatore della media. Grazie al fatto che $F_{pois(\lambda)}(k) = 1 - F_{\chi^2(2(k+1))}(2\lambda)$, l'intervallo di confidenza bilaterale a livello $1-\alpha$ è:

$$\left(\frac{1}{2n}F^{-1}_{\chi^2(2n\bar{X}_n)}\left(\frac{\alpha}{2}\right), \frac{1}{2n}F^{-1}_{\chi^2(2n\bar{X}_n+2)}\left(1-\frac{\alpha}{2}\right)\right)$$

Questo intervallo può essere approssimato a:

Intervallo bilaterale

$$\left(\bar{X}_n - \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{\bar{X}_n}{n}}, \bar{X}_n + \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{\bar{X}_n}{n}}\right)$$

Intervallo unilaterale sinistro

$$\left(0, \bar{X}_n + \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{\bar{X}_n}{n}}\right)$$

Intervallo unilaterale destro

$$\left(\bar{X}_n - \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{\bar{X}_n}{n}}, +\infty\right)$$

Larghezza

$$l = 2\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{\bar{X}_n}{n}} \Rightarrow n = \left\lceil \frac{4\left(\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\right)^2 \bar{X}_n}{l^2} \right\rceil$$