Guida pazza di Probabilitá

Riccardo Graziani

Anno Accademico 2024/2025

Contents

1 Introduzione

Questa guida copre la maggior parte dei tipi di esercizi che vengono proposti all'esame di Probabilitá e Statistica, elencando casistiche, istruzioni passo passo per la risoluzione ed esempi pratici.

2 Notazione di base

Di seguito sono elencate **tutte** le notazioni usate in questo documento.
(Def. *Distribuzione Normale*) Simbolo $N(\mu, \sigma^2)$, ossia media μ e varianza σ^2,

presenta la seguente densitá.

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp(-\frac{(x-\mu)^2}{2\sigma^2})$$

(Def. *Distribuzione Bernoulli*) Simbolo $Bern(p)$, una variabile di questa distribuzione assume solo due valori:

$$P(X = 1) = p, \quad P(X = 0) = 1 - p, \quad p \in [0, 1]$$

e si hanno i seguenti valori notevoli di $E[X]$ e $Var(X)$:

$$E[X] = p, \quad Var(X) = p(1 - p)$$

(Def. *Distribuzione Rademacher*) Simbolo $Rad(p)$, una variabile di questa distribuzione assume solo due valori:

$$P(X = 1) = p, \quad P(X = -1) = 1 - p, \quad p \in [0, 1]$$

e si hanno i seguenti valori notevoli di $E[X]$ e $Var(X)$:

$$E[X] = 2p - 1, \quad Var(X) = 4p(1 - p), \quad p \neq \frac{1}{2}$$

$$E[X] = 0, \quad Var(X) = 1, \quad p = \frac{1}{2}$$

(Def. *Media*) Simbolo $E[X]$ si calcola facendo la sommatoria del valore di X_i per la probabilitá associata p_i.

$$E[X] = \sum_i x_i p_i$$

(Def. *Varianza*) Simbolo $Var(X)$ si calcola facendo la differenza tra la media dei **valori** della variabile al **quadrato** e il **quadrato** della media.

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

(Def. *Covarianza*) Simbolo $Cov(X, Y)$, si calcola secondo la seguente formula.

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

(Def. *Densitá*) Simbolo $f(x)$ viene calcolata come la derivata della funzione $F_X(x)$.

$$f(x) = \frac{d}{dx} F_X(x)$$

(Def. *Densitá semplice*) Simbolo $f(x)$ viene calcolata come $(b-a)^{-1}$ dove a, b sono gli estremi dello spazio uniforme.

$$f(x) = (b-a)^{-1} = \frac{1}{b-a}, \quad x \in [a, b]$$

(Def. *Media Uniforme*) Simbolo $E[X]$ si calcola facendo la somma tra gli **estremi** del campo continuo diviso per 2.

$$E[X] = \int_a^b x \cdot f(x) dx = \frac{a+b}{2}, \quad [a, b]$$

(Def. *Varianza Uniforme*) Simbolo $Var(X)$ si calcola come il **quadrato** della differenza tra gli **estremi** del campo continuo diviso per 12.

$$Var(X) = \int_a^b x^2 \cdot f(x) dx - E[X]^2 = \frac{(b-a)^2}{12}$$

(Def. *Densitá esponenziale)* Simbolo $f_Y(y)$ si calcola secondo la seguente formula, dove il parametro λ indica il grado dell'esponenziale.

$$f_Y(y) = \lambda e^{-\lambda y}, \quad y \geq 0$$

(Def. *Media esponenziale*) Simbolo $E[Y]$ si calcola come l'inverso di λ.

$$E[Y] = \frac{1}{\lambda}$$

(Def. *Varianza epsonenziale*) Simbolo $Var(Y)$ si calcola come il **quadrato** dell'inverso di λ.

$$Var(Y) = \frac{1}{\lambda^2}$$

3 Calcolo media e varianza

3.1 Variabile discreta

Se il vostro testo recita qualcosa come: "X é una variabile aleatoria discreta tale che $P(X = N)$", dove N é un numero, allora seguite i passaggi qui sotto. Passaggi di risoluzione (per le definizioni riferire **Notazione di Base**):

- **Verificare** che le probabilitá sommino ad 1.
- Calcolare la **Media** $E[X]$.
- Calcolare la **Varianza** $Var(X)$.

Esempio concreto:

$$P(X=-4)=\frac{1}{6}, P(X=-3)=\frac{1}{3}, P(X=1)=\frac{1}{6}, P(X=3)=\frac{1}{3}$$

$$P(X)=\sum_i p_i=\frac{1}{6}+\frac{1}{3}+\frac{1}{6}+\frac{1}{3}=\frac{1+2+1+2}{6}=1$$

$$E[X]=\sum_i x_i p_i=-4\cdot\frac{1}{6}-3\cdot\frac{1}{3}+1\cdot\frac{1}{6}+3\cdot\frac{1}{3}=-\frac{1}{2}$$

$$E[X^2]=(-4)^2\cdot\frac{1}{6}-3^2\cdot\frac{1}{3}+1^2\cdot\frac{1}{6}+3^2\cdot\frac{1}{3}=\frac{53}{6}$$

$$Var(X)=E[X^2]-E[X]^2=\frac{53}{6}-(\frac{1}{2})^2=\frac{103}{12}$$

3.2 Variabile uniforme

Se il vostro testo recita qualcosa come: "$X \doteq Y$ é una variabile aleatoria uniforme continua su $[a,b]$", con a,b numeri, allora seguite i passaggi qui sotto. Passaggi di risoluzione (per le definizioni riferire **Notazione di Base**):

- Calcolare la **Media Uniforme** $E[X]$.
- Calcolare la **Varianza Uniforme** $Var(X)$.

Esempio concreto (caso in cui X uniforme semplice):

$$X=Y\sim Unit(-2,-1), \quad a<b$$

$$E[Y]=\frac{b+a}{2}=\frac{-3}{2}, \quad Var(Y)=\frac{(b-a)^2}{12}=\frac{1}{12}$$

Esempio concreto (caso in cui X uniforme come funzione):

$$X=Z^3, \quad Z\sim Unit(-2,2), \quad a<b$$

$$E[X]=E[Z^3]=\int_{-2}^{2} z^3\cdot\frac{1}{4}dz=\frac{1}{4}\cdot\int_{-2}^{2} z^3 dz=0$$

$$E[X^2]=E[Z^6]=\frac{1}{4}\cdot\int_{-2}^{2} z^6 dz=\frac{z^7}{7}|_{-2}^{2}=\frac{128}{7}+\frac{128}{7}=\frac{1}{4}\cdot\frac{256}{7}=\frac{64}{7}$$

$$Var(X)=E[X^2]-E[X]^2=\frac{64}{7}-0^2=\frac{64}{7}$$

3.3 Funzione di ripartizione

Se il vostro testo recita qualcosa come: "X ha funzione di ripartizione $F_X(x)$ data da $F_X(x) = \dots,\ x \in \mathbb{R}$", allora seguite i passaggi qui sotto.
Passaggi di risoluzione (per le definizioni riferire **Notazione di Base**):

- Calcolare la **densitá** di $F_X(x)$.
- Calcolare la **Media** della **densitá** come nella media uniforme.
- Calcolare la **Varianza**.

Esempio concreto:

$$F_X(x) = sin(x) \cdot 1_{[0,\frac{\pi}{2})}(x) + 1_{[\frac{\pi}{2},\infty)}(x)$$

$$F_X(x) = \begin{cases} 0 \text{ se } x < 0 \\ sin(x) \text{ per } x \in [0, \frac{\pi}{2}) \\ 1 \text{ per } x \geq \frac{\pi}{2} \end{cases} \quad , \quad f_X(x) = \frac{d}{dx}F_X(x) = cos(x) \cdot 1_{[0,\frac{\pi}{2})}$$

$$E[X] = \int x \cdot cos(x)dx = xsin(x) + cos(x) \Rightarrow$$

$$E[X] = [xsin(x) + cos(x)]_0^{\frac{\pi}{2}} = \frac{\pi}{2} \cdot 1 + cos(\frac{\pi}{2}) - (0 + cos(0)) = \frac{\pi}{2} - 1$$

$$E[X^2] = \int x^2 \cdot cos(x)dx = [x^2 sin(x) + 2xcos(x) - 2sin(x)]_0^{\frac{\pi}{2}} = \frac{\pi^2}{4} \cdot 1 - 2 - 0 = \frac{\pi^2}{4} - 2$$

$$Var(X) = E[X^2] - E[X]^2 = \frac{\pi^2}{4} - 2 - (\frac{\pi}{2} - 1)^2 = \frac{\pi^2}{4} - 2 - \frac{\pi^2}{4} + \pi - 1 = \pi - 3$$

3.4 Trasformazioni

Se il vostro esercizio non combacia con nessuno dei precedenti, siete entrati nel territorio delle **trasformazioni** di variabili aleatorie. Di seguito sono elencati i procedimenti risolutivi per i piú comuni tipi di trasformazioni.

3.4.1 Lineari

Si presentano nella forma $X = aY + b$. Il procedimento per la risoluzione é semplice ma richiede le seguenti formule specifiche:

- **Media**: $E[X] = aE[Y] + b$.
- **Varianza**: $Var(X) = a^2 Var(Y)$.

Esempio concreto (con **Distribuzione Normale**):

$$Y \sim N(2, 1), \quad X = 3Y - 5, \quad E[Y] = 2, Var(Y) = 1$$

$$E[X] = 3 \cdot E[Y] - 5 = 6 - 5 = 1, \quad Var(X) = 3^2 Var(Y) = 9$$

3.4.2 Monotone invertibili

Si presentano nella forma $X = g(Y)$ dove g é una funzione **monotona** e **differenziabile** (esempio comune $Y \sim exp(\lambda = n)$).
Il procedimento di risoluzione é il seguente:

- (**Opzionale**): calcola $f_X(x)$ a partire da $f_Y(y)$ secondo la formula

$$f_X(x) = f_Y(g^{-1}(x)) \cdot |\frac{d}{dx} g^{-1}(x)|$$

  e decidi chi tra le due densitá é piú facile da **integrare**.

- Calcola la **Media** come:

$$E[X] = \int x f_{(X/Y)}(x/y) d(x/y)$$

  scegliendo come estremi di integrazione il **supporto** della variabile Y (ad esempio per exp vale $Y > 0$).

- Calcola la **Varianza** come:

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

Esempio concreto (usiamo la exp):

$$X = e^Y, \quad Y \sim Exp(\lambda = 4), \quad f_Y(y) = 4e^{-4y}, \quad y > 0$$

$$E[X] = E[e^Y] = \int_0^\infty e^y \cdot 4e^{-4y} dy = 4 \int_0^\infty e^{-3y} dy = 4[-\frac{1}{3} e^{-3y}]_0^\infty = \frac{4}{3}$$

$$E[X^2] = E[e^{2Y}] = \int_0^\infty e^{2y} \cdot 4e^{-4y} dy = 4 \int_0^\infty e^{-2y} dy = 4[-\frac{1}{2} e^{-2y}]_0^\infty = \frac{4}{2} = 2$$

$$Var(X) = E[X^2] - E[X]^2 = 2 - \frac{16}{9} = \frac{2}{9}$$

3.4.3 Potenza

Si presentano nella forma $X = Y^3$ e $Y \sim Unif(-1, 1)$.
Il procedimento per la risoluzione é il seguente:

- Calcolo la **Media** come:

$$E[X] = E[Y^n] = \int y^n f_Y(y) dy$$

  con un trucco: quando Y ha densità **simmetrica** su $(-a, a)$ e n é **dispari** allora $E[Y^n] = 0$. Altrimenti se $Y \sim N(0, 1)$ possiamo usare i **momenti noti**: $E[Y^2] = 1, E[Y^4] = 6, E[Y^6] = 15, \ldots$

- Calcolo la **Varianza** come:

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

Esempio concreto (vengono usate le **formule dei momenti** della Normale per semplificare i calcoli):

$$X = Y^2, \quad Y \sim N(0,1), \quad f_X(x) = \frac{1}{\sqrt{2\pi 1^2}} exp(-\frac{(y)^2}{2 \cdot 1^2}) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}, \quad y \in \mathbb{R}$$

$$E[X] = E[Y^2] = \int_{-\infty}^{\infty} y^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = 1$$

$$E[X^2] = E[Y^4] = \int_{-\infty}^{\infty} y^4 \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = 3$$

$$Var(X) = E[X^2] - E[X]^2 = 3 - 1^2 = 2$$

3.4.4 Composizione

Nel caso in cui ci sia una combinazione delle precedenti trasformazioni, semplicemente scomporre in passi la risoluzione e riferirsi alle trasformazioni una ad una.

4 Variabili correlate

In questo esercizio vengono date tre variabili ξ_1, ξ_2, ξ_3 indipendenti e identicamente distribuite con una delle distribuzioni notevoli. Vengono poi definite due variabili aleatorie X, Y come una combinazione delle tre variabili iniziali. Vengono proposti esempi di risoluzione per le distribuzioni piú comuni nell'esame. Il procedimento per la risoluzione (nel caso di **Rademacher** e **Bernoulli**) é il seguente:

- Identificare la **distribuzione** usata e scriversi il valore di $E[\xi_i]$ e $Var(\xi_i)$ sfruttando le conoscenze sulla distribuzione.
- Calcolare $E[X]$ e $Var(X)$ seguendo la definizione data.
- Calcolare $E[Y]$ e $Var(Y)$ seguendo la definizione data.
- Calcolare $Cov(X, Y)$, che si traduce in calcolare $E[XY]$ e applicare la formula della $Cov(X, Y)$ Se $Cov(X, Y) \neq 0$ allora sono **dipendenti**, altrimenti va deciso dopo aver svolto il passo successivo.
- Creare la **tabella** della legge congiunta prendendo i valori possibili di X, Y a partire da tutte le combinazioni di ξ_1, ξ_2, ξ_3 in $[0, 1]$ se Bernoulli o $[-1, 1]$ se Rademacher. Per verificare l'indipendenza bisogna che, per ogni coppia (x, y) si abbia $P(X = x, Y = y)$.

4.1 Distribuzione Bernoulli

$$\xi_1, \xi_2, \xi_3 \in Bern(\frac{1}{2}), \quad X = \xi_1 - \xi_2, \quad Y = \xi_3 \cdot (\xi_1 + \xi_2)$$

Della distribuzione di Bernoulli sappiamo che:

$$P(\xi_i = 1) = P(\xi_i = 0) = \frac{1}{2}, \quad E[\xi_i] = \frac{1}{2}, \quad Var(\xi_i) = \frac{1}{2}(1 - \frac{1}{2}) = \frac{1}{4}$$

Calcoliamo media e varianza di X, Y (i quadrati sono irrilevanti poiché lavoriamo solo con $0, 1$):

$$E[X] = E[\xi_1] - E[\xi_2] = 0, \quad E[X^2] = E[(\xi_1 - \xi_2)^2] = E[\xi_1^2 - 2\xi_1\xi_2 + \xi_2^2] = \frac{1}{2} - 2\frac{1}{4} + \frac{1}{2} = \frac{1}{2}$$

$$Var(X) = E[X^2] - E[X]^2 = \frac{1}{2}$$

$$E[Y] = E[\xi_3] \cdot E[\xi_1 + \xi_2] = \frac{1}{2} \cdot 1 = \frac{1}{2}, \quad E[Y^2] = E[\xi_3^2(\xi_1^2 + 2\xi_1\xi_2 + \xi_2^2)] = \frac{1}{2} \cdot \frac{3}{2} = \frac{3}{4}$$

$$Var(Y) = E[Y^2] - E[Y]^2 = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$$

Calcoliamo la covarianza di X, Y:

$$E[XY] = E[(\xi_1 - \xi_2) \cdot \xi_3 \cdot (\xi_1 + \xi_2)] = E[\xi_3(\xi_1^2 - \xi_2^2)] = \frac{1}{2} \cdot (\frac{1}{2} - \frac{1}{2}) = 0$$

$$Cov(X, Y) = E[XY] - E[X]E[Y] = 0 - 0 \cdot \frac{1}{2} = 0$$

Non possiamo ancora decidere se X, Y sono indipendenti o meno, quindi andiamo avanti. Creiamo la tabella della legge congiunta per X, Y:

$$\xi_1, \xi_2, \xi_3 \in [0, 1] \Rightarrow 2^3 \text{ combinazioni}$$

ξ_1	ξ_2	ξ_3	$X = \xi_1 - \xi_2$	$Y = \xi_3 \cdot (\xi_1 + \xi_2)$
0	0	0	0	0
0	0	1	0	0
0	1	0	-1	0
0	1	1	-1	1
1	0	0	1	0
1	0	1	1	1
1	1	0	0	0
1	1	1	0	2

Table 1: Tutte le combinazioni possibili di (ξ_1, ξ_2, ξ_3) con i corrispondenti valori di X e Y

Per la legge congiunta, contare quante volte compare $P(X = x, Y = y)$ nella tabella:

$Y\backslash X$	-1	0	1
0	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{8}$
1	$\frac{1}{8}$	0	$\frac{1}{8}$
2	0	$\frac{1}{8}$	0

Table 2: Legge congiunta di $X = \xi_1 - \xi_2$ e $Y = \xi_3(\xi_1 + \xi_2)$

Ora possiamo determinare se X, Y sono indipendenti o meno calcolando i marginali per ogni probabilitá:

$$P(X = x) \Rightarrow P(X = -1) = \frac{1}{4}, \quad P(X = 0) = \frac{1}{2}, \quad P(X = 1) = \frac{1}{4}$$

$$P(Y = y) \Rightarrow P(Y = -0) = \frac{5}{8}, \quad P(X = 1) = \frac{1}{4}, \quad P(X = 2) = \frac{1}{8}$$

E controlliamo per ogni coppia (x, y) se il prodotto dei marginali combacia con il valore presente nella legge congiunta, ad esempio:

$$P(X = 0, Y = 0) = \frac{3}{8} \Rightarrow P(X = 0) = \frac{1}{2}, \quad P(Y = 0) = \frac{5}{8}$$

$$P(X = 0) \cdot P(Y = 0) = \frac{5}{16} \neq \frac{3}{8}$$

Quindi concludiamo che X, Y sono dipendenti.

4.2 Distribuzione Rademacher

$$\xi_1, \xi_2, \xi_3 \in Rad(\frac{1}{2}), \quad X = \xi_1 \cdot \xi_2, \quad Y = \xi_1 \cdot (\xi_2 + \xi_3)$$

Dalla distribuzione di Rademacher sappiamo che:

$$P(\xi_i = 1) = P(\xi_i = -1) = \frac{1}{2}, \quad E[\xi_i] = 0, \quad Var(\xi_i) = 1$$

Calcoliamo media e varianza di X, Y (tenendo a mente quando abbiamo il -1):

$$E[X] = E[\xi_1 \cdot \xi_2] = E[\xi_1] \cdot E[\xi_2] = 0 \cdot 0 = 0$$

$$Var(X) = E[X^2] - E[X]^2 = E[(\xi_1 \cdot \xi_2)^2] - 0 = 1^2 \cdot 1^2 = 1$$

$$E[Y] = E[\xi_1 \cdot (\xi_2 + \xi_3)] = E[\xi_1] \cdot (E[\xi_2] - E[\xi_3]) = 0 \cdot (0 - 0) = 0$$

$$E[Y^2] = E[\xi_1^2 \cdot (\xi_2^2 - 2\xi_2\xi_3 + \xi_3^2)] = E[\xi_2^2 - 2\xi_2\xi_3 + \xi_3^2] = 1 + 1 = 2\cdot$$

$$Var(Y) = E[Y^2] - E[Y]^2 = 2 - 0^2 = 2$$

Calcoliamo la covarianza di X, Y:

$$E[XY] = E[(\xi_1 \cdot \xi_2)[\xi_1(\xi_2 - \xi_3)]] = E[\xi_1^2 \cdot \xi_2(\xi_2 - \xi_3)] = E[\xi_2(\xi_2 - \xi_3)] \Rightarrow$$

$$E[\xi_2^2 - \xi_2\xi_3] = 1 - E[\xi_2\xi_3] = 1 - 0 \cdot 0 = 1$$

$$Cov(X,Y) = E[XY] - E[X]E[Y] = E[XY] - 0 \cdot 0 = E[XY] = 1 \neq 0$$

Quindi poiché $Cov(X,Y) \neq 0$ possiamo affermare che X, Y non sono indipendenti. Infine calcoliamo la tabella della legge congiunta:

ξ_1	ξ_2	ξ_3	$X = \xi_1 - \xi_2$	$Y = \xi_3 \cdot (\xi_1 + \xi_2)$
-1	-1	-1	1	0
-1	-1	1	1	2
-1	1	-1	-1	-2
-1	1	1	-1	0
1	-1	-1	-1	0
1	-1	1	-1	2
1	1	-1	1	2
1	1	1	1	0

Table 3: Tutte le combinazioni possibili di (ξ_1, ξ_2, ξ_3) con i corrispondenti valori di X e Y

Per la legge congiunta, contare quante volte compare P (X = x, Y = y) nella tabella:

$Y \backslash X$	-1	1
-2	$\frac{2}{8}$	0
0	$\frac{2}{8}$	$\frac{2}{8}$
2	0	$\frac{2}{8}$

Table 4: Legge congiunta di $X = \xi_1 - \xi_2$ e $Y = \xi_3(\xi_1 + \xi_2)$

4.3 Distribuzione Esponenziale

5 Le tre stime

In questo esercizio abbiamo sempre: $X_N, 800 \leq N \leq 1200$ variabili aleatorie indipendenti ed identicamente distribuite con una delle distribuzioni comuni (es. Bernoulli o Binomiale) di parametro p, $\frac{1}{500} \leq p \leq \frac{1}{200}$ e una probabilitá P alta, ossia $0.95 - 0.99$. Viene richiesto di dare una stima tramite: disuguaglianza di **Chebyshev**, approssimazione di **Poisson**, approssimazione **normale**. Il procedimento per la risoluzione é il seguente:

- Calcolare Media $E[S]$ come: $N \cdot p$ e Varianza $Var(S)$ come: $N \cdot p \cdot (1-p)$.
- Per calcolare **Chebyshev** vogliamo trovare M tale che $P(S < M) \leq 1-P$. Preso $E[S]$, possiamo dire che:

$$P(S < M) = P(S - E[S] > M - E[S]) \leq \frac{(Var[S])}{(M - E[S])^2}$$

$$\frac{(Var[S])}{(M-E[S])^2} \leq 1-P \Rightarrow (M-E[S])^2 \geq \frac{Var[S]}{1-P}$$

e infine risolvere per M e approssimare ad un numero naturale.

- Per calcolare **Poisson** usare: $\lambda = N \cdot p$ e cercare nella tavola di Poisson il primo valore k per cui vale $F_{Poisson(\lambda)}(k) \geq P$ (vedere come va letta la tavola).

- Per calcolare la **Normale** possiamo approssimare $S \sim Bin(800, \frac{1}{400})$ come una distribuzione normale $N(\mu, \sigma^2)$ con $\mu = E[S]$ e $\sigma^2 = Var[S]$. Vogliamo trovare N tale che:

  $$P(S \leq N) \approx P(\frac{S-\mu}{\sigma} \leq \frac{N+0.5-\mu}{\sigma}) = \Phi(\frac{N+0.5-\mu}{\sigma}) \geq P$$

  Cercare nella tavola della Normale (vedere come va letta la tavola) il valore di $\Phi(k)$ tale che $\Phi(k) \approx P$ e poi risolvere:

  $$(\frac{N+0.5-\mu}{\sigma}) \geq k$$

  in funzione di N sempre avendo cura di arrotondare N ad un intero.

Esempio pratico:

$$X_1, \dots, X_{1000}, \quad X_i \sim Bern(\frac{1}{500}), \quad S = \sum_{i=1}^{1000} X_i$$

$$E[X_i] = p = \frac{1}{500}, \quad E[S] = N \cdot p = 1000 \cdot \frac{1}{500} = 2$$

$$Var[X_i] = p \cdot (1-p) = \frac{1}{500} \cdot \frac{499}{500} = \frac{499}{250000}, \quad Var[S] = N \cdot Var[S] = 1000 \cdot \frac{499}{250000} = \frac{499}{250} \approx 2$$

Dare una stima per $N_* = min\{k \in \mathbb{N} : P(S \leq k) \geq 0.98\}$ con Chebyshev, Poisson e Normale.

Chebyshev:

$$P(S < M) \leq 1-P \leq \frac{(Var[S])}{(M-E[S])^2} \Rightarrow \frac{2}{(M-2)^2} \leq 0.02$$

$$\frac{2}{(M-2)^2} \leq 0.02 \Rightarrow (M-2)^2 \geq \frac{2}{0.02} = (M-2) \geq \sqrt{100} = M \geq 12$$

Poisson, dati N grande e p piccolo, approssimiamo $S \sim Bin(1000, \frac{1}{500})$:

$$\lambda = N \cdot p = 1000 \cdot \frac{1}{500} = 2$$

Sulle tavole di Poisson cerchiamo dove $\lambda = 2$ e guardiamo il primo k tale che $k \geq P$, nel nostro caso troviamo $k = 5$.
Normale:

$$S \approx N(\mu = 2, \sigma^2 = 2), \quad \Phi(\frac{N + 0.5 - \mu}{\sigma}) \geq P$$

$$\Phi(\frac{N - 1.5}{\sqrt{2}}) \geq 0.98$$

Cerchiamo nella tavola il primissimo valore che supera la nostra probabilitá, in questo caso troviamo il valore 0.9803: da qui sommiamo il valore di z corrispondente sulla riga e sulla colonna, nel nostro caso abbiamo in riga $z = 2.0$ e in colonna $z = 0.06$, quindi il valore finale é $z = 2.06$. Ora risolviamo per N:

$$\frac{N - 1.5}{\sqrt{2}} \geq 2.06 \Rightarrow N \geq 3.56 \cdot \sqrt{2} \approx 5$$