

ESERCITAZIONE

Prof. Cerchiara Rocco Roberto

Materiale e Riferimenti

- 1. Lucidi distribuiti in aula**
- 2. Loss Models 1998 – First Edition**

1. Variabili condizionate e
Distribuzione di Poisson
Misturata
2. Distribuzione Lognormale
3. Stima del parametro della
Poisson Semplice con l'effetto
delle “esposizioni”

EVENTI CONDIZIONATI:

$$P(A | B) = \frac{P(A \cap B)}{P(B)}$$

che equivale a

$$P(A \cap B) = P(A | B) \cdot P(B)$$

se A e B sono indipendenti

$$P(A \cap B) = P(A) \cdot P(B)$$

quindi avremo

$$P(A | B) = \frac{P(A) \cdot P(B)}{P(B)} = P(A)$$

nel discreto:

$$P(X = x | Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}$$

nel continuo

$$f_{X|Y}(x | y)dx = \frac{f_{X,Y}(x, y)dxdy}{f_Y(y)dy}$$

$$E(X | Y) = \sum_x x P(X = x | Y = y)$$

$$E(X) = \sum_y E(X | Y = y) P(Y = y)$$

dim:

$$E(X) = \sum_y \sum_x x P(X = x | Y = y) P(Y = y)$$

$$E(X) = \sum_y \sum_x x P(X = x, Y = y) =$$

$$E(X) = \sum_x x \sum_y P(X = x, Y = y) =$$

$$E(X) = \sum_x x P(X = x)$$

POISSON MISTURATA (DAIKIN ET AL 1994)

$N \sim \text{Poisson}(\lambda)$ e

\tilde{q} variabile aleatoria *disturbo o misturante*,
distribuita con funzione di densità $f_{\tilde{q}}(q)$

$$p_k(\tilde{q}) = \int_0^{\infty} P(k | q) f_{\tilde{q}}(q) dq = \int_0^{\infty} \frac{e^{-\lambda q} (\lambda q)^k}{k!} f_{\tilde{q}}(q) dq$$

$$E(\tilde{q}) = 1$$

Proprietà

- $E(N) = \lambda$
- $Var[N] = \lambda + \lambda^2 Var[\tilde{q}]$

La varianza della misturata è pari alla varianza dovuta alla poissoniana – uguale alla media e cioè λ - più quella aggiunta dalla distribuzione misturante: $\lambda^2 Var[\tilde{q}]$.

OPPURE: POISSON MISTURATA (BÜHLMANN)

$N \sim \text{Poisson}(\Lambda)$ e

Λ variabile aleatoria *misturante*, distribuita
con funzione di densità $f_{\Lambda}(\lambda)$

$$p_k(\Lambda) = \int_0^{\infty} P(k | \lambda) f_{\Lambda}(\lambda) d\lambda = \int_0^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} f_{\Lambda}(\lambda) d\lambda$$

Proprietà

- $E(N) = E(\Lambda)$
- $Var[N] = E[\Lambda] + Var[\Lambda]$

La varianza della mistura è pari alla varianza dovuta alla poissoniana – uguale alla media e cioè $E[\Lambda]$ - più quella aggiunta dalla distribuzione misturante: $Var[\Lambda]$.

CASO POISSON – GAMMA

Se si usa come funzione misurante una *Gamma*(α, τ):

$$f_{\Lambda}(\lambda) = \frac{\tau^{\alpha} e^{-\lambda\tau} \lambda^{\alpha-1}}{\Gamma(\alpha)} \quad \text{con } \alpha, \tau > 0$$

$$\text{e } \Gamma(\alpha) = \int_0^{\infty} e^{-t} t^{\alpha-1} dt$$

$$E[\Lambda] = \frac{\alpha}{\tau} \quad \dots \quad \text{Var}[\Lambda] = \frac{\alpha}{\tau^2}$$

si ottiene per N una distribuzione *Binomiale negativa* (τ, α):

$$p_k(\alpha, \tau) = \binom{k + \alpha - 1}{\alpha - 1} \left(\frac{\tau}{1 + \tau} \right)^{\alpha} \left(1 - \frac{\tau}{1 + \tau} \right)^k$$

con

$$E[N] = \frac{\alpha}{\tau}$$

$$\text{Var}[N] = \frac{\alpha}{\tau} \left(1 + \frac{1}{\tau} \right) = E[N] + \frac{\alpha}{\tau^2}$$

La binomiale negativa converge ad una *Poisson* in particolari condizioni:

$$\text{BinNeg}(\alpha, \tau) \xrightarrow[\substack{\tau \rightarrow 0 \\ \alpha/\tau = \text{cost.}}]{\alpha \rightarrow \infty} \text{Poisson}\left(\frac{\alpha}{\tau}\right).$$

Quindi il parametro α può essere interpretato come una misura del grado di eterogeneità del portafoglio; quando α è grande, il portafoglio è abbastanza omogeneo, i sinistri accadono in maniera abbastanza casuale e la loro distribuzione si avvicina a quella di *Poisson*.

Esercizio 1

Si è rilevata la seguente distribuzione delle polizze secondo il numero di sinistri.

Numero sinistri	Numero polizze
0	3288
1	642
2	66
3	4
totale	4000

E' ragionevole modellizzare tale distribuzione con una Poissoniana? Con quali parametri?

Soluzione esercizio 1

Numero sinistri	Numero polizze	Frequenze relative	Poisson (0,1965)	Numero polizze stimato
0	3288	0,8220	0,82160	3286
1	642	0,1605	0,16144	646
2	66	0,0165	0,01586	63
3	4	0,0010	0,00104	4
totale	4000	1	circa 1	c. 4000

$$\bar{X} = 0.1965 \quad \text{AND} \quad S^2 = 0.196888$$

stima con il metodo dei momenti:

$$H_p : \hat{\lambda} = \bar{X} \Rightarrow \text{Poisson}(0.1965)$$

Esercizio 2

Si è rilevata la seguente distribuzione delle polizze secondo il numero di sinistri.

MERCATO RCAUTO ITALIA 1993.

Numero sinistri	Numero polizze
0	90550
1	8733
2	652
3	56
4	9
totale	10000

E' ragionevole modellizzare tale distribuzione con una Poissoniana? Con quali parametri?

Soluzione esercizio 2

$$\bar{X} = 0.10241 \quad \text{AND} \quad S^2 = 0.109402$$

Numero sinistri	Numero polizze	Frequenze relative	Poisson (0,10241)	Numero polizze stimato
0	90550	0,90550	0,90266	90266
1	8733	0,08733	0,09244	9244
2	652	0,00652	0,00473	473
3	56	0,00056	0,00016	16
4	9	0,00009	0,00000	0
<u>totale</u>	<u>10000</u>	<u>1</u>	<u>circa 1</u>	<u>c. 10000</u>

stima con il metodo dei momenti:

$$H_p : \hat{\lambda} = \bar{X} \Rightarrow \text{Poisson}(0.10241)$$

Soluzione esercizio 2 (Continua)

La media e la varianza empiriche sono sufficientemente diverse per giustificare una distribuzione del numero dei sinistri con 2 parametri.

Usando la Binomiale Negativa (Polya) e stimandone i parametri col metodo dei momenti si ottengono migliori risultati

Numero sinistri	Numero polizze	Frequenze relative	Stime Poisson	Stime Polya
0	90550	0,90550	90266	90568
1	8733	0,08733	9244	8682
2	652	0,00652	473	694
3	56	0,00056	16	52
4	9	0,00009	0	4
<u>totale</u>	<u>10000</u>	<u>1</u>	<u>circa 1</u>	<u>c. 10000</u>

ESERCIZIO 3

La variabile N numero dei sinistri sia distribuita secondo la tabella seguente:

Numero sinistri	Numero polizze
0	17353
1	1414
2	620
3	25
totale	19412

Si ipotizzi che l'ammontare Z di un sinistro sia distribuito secondo una esponenziale negativa di media 1000:

$$F_Z(z) = 1 - e^{-0.001z} \quad \dots \quad f_Z(z) = 0.001 \cdot e^{-0.001z}$$

Supponendo i costi dei sinistri indipendenti tra di loro e dal numero dei sinistri e posto

$$X = \sum_{i=0}^N Z_i,$$

si calcoli:

- $E(X|N=0)$, $E(X|N=1)$, $E(X|N=2)$, $E(X|N=3)$
- $E(X)$

SOLUZIONE

N	$P(N=n)$
0	89.39%
1	7.28%
2	3.19%
3	0.13%

$$E(X | N = 0) = 0$$

$$E(X | N = 1) = E(Z) = 1000$$

$$E(X | N = 2) = E(Z + Z) = 2E(Z) = 2000$$

$$E(X | N = 3) = E(Z + Z + Z) = 3E(Z) = 3000$$

$$E(X) = \sum_{i=0}^3 E(X | N = i) P(N = i) = 140.58$$

ESERCIZIO 4

Il numero di sinistri N per polizza RCAuto di una compagnia sia distribuito secondo Poisson(λ). La compagnia ritenga di avere un portafoglio omogeneo. Per l'anno successivo ci si attende:

- 1. una diminuzione generalizzata della frequenza sinistri pari al 5% (trend decrescente del mercato italiano);**
- 2. un fattore di disturbo non generalizzato (effetto patente a punti? Blocco del traffico?) sintetizzato da q :**

Valore	Probabilità
90%	0.4
110%	0.4
100%	0.2

Calcolare :

- 1. il numero medio di sinistri commessi da un guidatore nell'anno successivo**
- 2. la probabilità che tale guidatore commetta 0 sinistri.**

SOLUZIONE

Modellizzazione:

$$N = \text{Poisson}(\lambda)$$

inserisco il trend decrescente (costante)

$$N = \text{Poisson}\left(\lambda \frac{95}{100}\right)$$

disturbo aleatorio:

$$\tilde{q} = \begin{cases} 0.9 & p(q = 0.9) = 0.4 \\ 1.0 & p(q = 1.0) = 0.2 \\ 1.1 & p(q = 1.1) = 0.4 \end{cases} \quad \text{N.B: } E(\tilde{q}) = 1$$

N si distribuisce quindi

$$P(N = k) = e^{-\lambda \frac{95}{100} \tilde{q}} \frac{\left(\lambda \frac{95}{100} \tilde{q}\right)^k}{k!}$$

***Poisson misturata* (non è Poisson pura! q è aleatorio)**

PUNTO 1

Dato che

$$E(X) = \sum_y E(X | Y = y)P(Y = y)$$

nel nostro caso diventa:

$$E(N) = \sum_i E(N | q = q_i)P(q = q_i)$$

$$E(N) = E(N | q = 0.9)(0.4) + \\ + E(N | q = 1.0)(0.2) + E(N | q = 1.1)(0.4)$$

cioè

$$= \lambda \frac{95}{100} 0.9(0.4) + \lambda \frac{95}{100} 1.0(0.2) + \lambda \frac{95}{100} 1.1(0.4)$$

$$E(N) = \lambda \frac{95}{100} 1 = \lambda \frac{95}{100} E(q) = \lambda \frac{95}{100}$$

PUNTO 2

$$P(N = 2)$$

$$= P(N = 2 | q = 0.9)P(q = 0.9) +$$

$$+ P(N = 2 | q = 1.0)P(q = 1.0) +$$

$$+ P(N = 2 | q = 1.1)P(q = 1.1)$$

$$= e^{-\lambda \frac{95}{100} 0.9} \frac{\left(\lambda \frac{95}{100} 0.9 \right)^2}{k!} (0.9) + \dots$$

$$e^{-\lambda \frac{95}{100}} \frac{\left(\lambda \frac{95}{100} \right)^2}{2!} \cdot (e^{-0.9} P(0.9) + e^{-1.0} P(1.0) + e^{-1.1} P(1.1))$$

$$p_2 \left(\lambda \frac{95}{100} \right) \cdot E(e^{-q}) \neq p_2 \left(\lambda \frac{95}{100} \right) \cdot E(q)$$

ESERCIZIO 5

Il numero di sinistri N per polizza RCAuto di una compagnia sia distribuito secondo Poisson(λ). La compagnia ritenga di suddividere il proprio portafoglio in 2 sole categorie: *buoni guidatori* e *cattivi guidatori*, caratterizzati da frequenza sinistri media pari a 10% e 30% rispettivamente. Detta α la percentuale di cattivi guidatori all'interno del portafoglio, calcolare:

- 3. la probabilità che un *cattivo guidatore* commetta più di un sinistro;**
- 4. il numero medio di sinistri causati da un guidatore scelto a caso**
- 5. il livello massimo di α affinché la probabilità che un guidatore a caso non commetta sinistro sia non inferiore a 85%.**

SOLUZIONE

Punto 1

$N_c =$ numero sinistri cattivo guidatore

$N_c = N \mid$ guidatore è cattivo $=$ Poisson(0.3)

$P(N_c > 1) = 1 - P(N_c = 0) - P(N_c = 1) =$

$= 1 - 0.7408 - 0.2222 = 0.0369$

Punto 2

N.B. λ è variabile aleatoria *BERNOULLIANA*:

$$\lambda = \begin{cases} 0,3 & p = \alpha \\ 0,1 & p = 1 - \alpha \end{cases}$$

$$\begin{aligned} P(N = k) &= P\{(N = k) \cap [(\lambda = 0.3) \cup (\lambda = 0.1)]\} = \\ &= P\{(N = k) \cap (\lambda = 0.3)\} + P\{(N = k) \cap (\lambda = 0.1)\} = \\ &= P(N = k | \lambda = 0.3) \cdot P(\lambda = 0.3) + \\ &\quad + P(N = k | \lambda = 0.1) \cdot P(\lambda = 0.1) = \\ &= P(N = k | 0.3) \cdot \alpha + P(N = k | 0.1) \cdot (1 - \alpha) \\ E(N) &= \sum_k k \cdot p_k(0.3) \cdot \alpha + k \cdot p_k(0.1) \cdot (1 - \alpha) \\ &= \sum_k k \cdot p_k(0.3) \cdot \alpha + \sum_k k \cdot p_k(0.1) \cdot (1 - \alpha) = \\ &= \alpha \cdot 0.3 + (1 - \alpha) \cdot 0.1 = \underline{\underline{E(\lambda)}} \end{aligned}$$

Punto 3

$$P(N = 0) \geq 0.85$$

$$p_0(0.3) \cdot \alpha + p_0(0.1) \cdot (1 - \alpha) \geq 0.85$$

$$0.741 \cdot \alpha + 0.905 \cdot (1 - \alpha) \geq 0.85$$

$$-0.164 \cdot \alpha \geq -0.058 \rightarrow \alpha \leq 33\%$$

ESERCIZIO 6

Dalle statistiche dei sinistri di una compagnia di assicurazioni si può assumere che se X è il costo del sinistro, allora $Y=\ln X$ è distribuita come Normale di media 6,012 e varianza 1,792.

1. Qual è la probabilità che un sinistro sia maggiore di 1200?
2. Qual è la probabilità che un sinistro sia compreso tra 0 e 200?
3. Qual è la probabilità che un sinistro sia compreso tra 200 e 500?

soluzioni

$$1 = 0,210$$

$$2 = 0,297$$

$$3 = 0,263$$

ESERCIZIO 7

Determinare la stima di massima verosimiglianza del parametro della Poisson Semplice per i seguenti dati

Anno	Esposizioni	Sinistri
1986	2.145	207
1987	2.452	227
1988	3.112	341
1989	3.458	335
1990	3.698	362
1991	3.872	359

Indichiamo con λ il parametro della Poisson per una singola esposizione.

Nell'anno k avremo e_k esposizioni.

Quindi la v.a. numero sinistri segue una distribuzione di Poisson con parametro λe_k .

Indicando con n_k il numero di sinistri in un anno allora la funzione di verosimiglianza è:

$$L = \prod_{k=1}^6 \frac{e^{-\lambda \cdot e_k} (\lambda \cdot e_k)^{n_k}}{n_k!}$$

da cui si ottiene

$$l = \text{Log}L = \sum_k \left[-\lambda \cdot e_k + n_k \log(\lambda \cdot e_k) - \log(n_k!) \right]$$

e quindi

$$dl / d\lambda = \sum_k \left[-e_k + n_k \lambda^{-1} \right] = 0$$

e

$$\hat{\lambda} = \frac{\sum_{k=1}^6 n_k}{\sum_{k=1}^6 e_k} = \frac{1.831}{18.737} = 0,09772$$

Al fine di verificare la bontà di adattamento di tale modello potremmo ricorrere al test del chi-quadro.

Se ipotizziamo che per ciascun anno il numero sinistri è il risultato della somma di un numero (esposizione) di v.a. iid, → possiamo applicare il Teorema del Limite Centrale (ipotesi Dist. Normale)

Possiamo usare il test statistico

$$Q = \sum_k \frac{(n_k - E_k)^2}{V_k},$$

dove E_k (val. atteso) = λe_k e V_k (varianza) = λe_k

Tale Q ha un distribuzione (approssimata) chi-quadro con gradi di libertà uguale al numero di punti meno il numero di parametri da stimare (5).