Probabilità III

Variabili casuali continue

- Definizioni principali.
- Valore atteso e Varianza.
- V.C. Notevoli: uniforme, gaussiana, X^2
- Teorema del limite centrale.
- Principio ristretto di equivalenza in variabilità.
- Convergenza in legge.

Variabile casuale continua: osservazioni

- *Osservazione:* una v.c. continua può assumere infinite osservazioni (o realizzazioni).
- *Conseguenza*: se ogni osservazione fosse equiprobabile la probabilità di una realizzazione sarebbe infinitesima.

$$P(X=x_i)\approx 0$$

- Per v. c. continue si preferisce definire la probabilità che la realizzazione cada in un intervallo.

$$P(x_{inf}\leq X\leq x_{sup})$$

- *Osservazione*:se i # reali sono tanti, gli intervalli sono di più.
- *Conseguenza:* difficile definire una v.c. continua mediante le probabilità, meglio un suo "parente stretto" (da cui ricavare le probabilità che interessano).

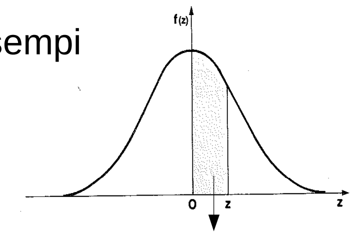
Variabile casuale continua X

Definita univocamente dalla densità di probabilità (d.d.p).

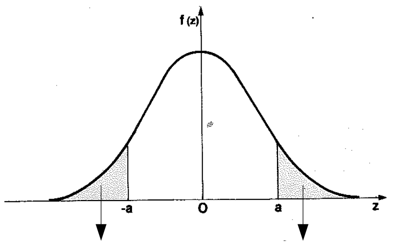
- <u>Densità di probabilità</u> $f(x)$: funzione reale di numeri reali la cui area sottesa fra due punti a e b, corrisponde alla probabilità che la v.c. X sia compresa fra a e b.

$$f:\mathbb{R}\rightarrow\mathbb{R} \qquad \int_a^b f(x)\,dx=P(a\leqslant X\leqslant b)$$

- Esempi



$P(0\leqslant Z\leqslant z)$



$P(Z<-a)$ $P(Z>a)$

$P(Z\leqslant -a)$ $P(Z\geq a)$

Densità di probabilità: proprietà

- L'area sottesa da tutta la curva deve essere unitaria

$$P(-\infty<X<+\infty)=1\Rightarrow\int_{-\infty}^{+\infty} f(x)\,dx=1$$

- Se è continua, la d.d.p. deve essere sempre positiva

$$f(x)\geq 0 \qquad \forall\, x\in\mathbb{R}$$

  - Dimostrazione per assurdo.
    Supponiamo che esista un intervallo $[a\,;b]$ in cui la d.d.p. è negativa. Allora si avrebbe

$$\int_a^b f(x)\,dx<0 \quad\Rightarrow\quad P(a\leqslant X\leqslant b)<0$$

Valore atteso E[.] e Varianza Var[.]

- Anche per le v.c. Continue valgono gli stessi concetti indicati da valore atteso e varianza.
- Valore atteso $E[X]$
  - v.c. discreta $E[X]=\sum_{i=1}^{M} p(x_i)x_i$
  - v.c. continua $E[X]=\int_{-\infty}^{+\infty} f(x)x\,dx$
- Varianza $Var[X]$
  - v.c. discreta $Var[X]=\sum_{i=1}^{M} p(x_i)(x_i-E[X])^2$
  - v.c. continua $Var[X]=\int_{-\infty}^{+\infty} f(x)(x-E[X])^2\,dx$
- *Osservazione:* nel calcolo si sostituisce la sommatoria con l'integrale e la distribuzione di probabilità con la densità di probabilità.

E[.] e Var[.]: proprietà

- Le proprietà del valore atteso e della varianza restano invariate anche per le v.c. continue.
- Variabili affini: $Y=aX+b$
  - $E[Y]=aE[X]+b$
  - $Var[Y]=a^2 Var[X]$
- Combinazione lineare di vv. cc. indipendenti $Y=\sum_{i=1}^{K} c_i X_i$
  - $E[Y]=\sum_{i=1}^{K} c_i E[X_i]$
  - $Var[Y]=\sum_{i=1}^{K} c_i^2 Var[X_i]$
- Teorema di Bienaymé - Čebičev

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

V.c. uniforme

- V.c. continua X le cui modalità appartengono all'intervallo $I=[a;b]$.

$$P(X<a)=P(X>b)=0 \quad\Leftrightarrow\quad f(x)=0 \quad \forall\, x\notin I$$

- Tutte le modalità sono equiprobabili

$$f(x)=k \quad \forall\, x\in I$$

  - calcoliamo l'integrale fra a e b

$$\int_a^b f(x)\,dx=\int_a^b k\,dx=(b-a)k$$

  - l'area sottesa da una d.d.p. è sempre unitaria

$$\int_a^b f(x)\,dx=1 \quad\Leftrightarrow\quad (b-a)k=1 \quad\Leftrightarrow\quad k=\frac{1}{b-a}$$

- Noti gli estremi la d.d.p è definita.
  - Si dice che

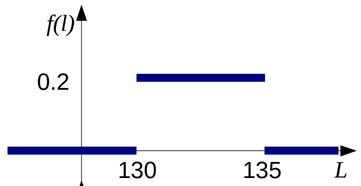
$$X\sim Unif(a,b)$$

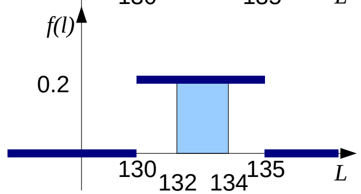
V.c. uniforme: esempio

Una macchina produce barre di acciaio la cui lunghezza oscilla fra 130 e 135 cm con egual probabilità. Calcolare la probabilità che si produca una barra compresa fra 132 cm e 134 cm

- La v.c. L: lunghezza della barra prodotta
  - Si ha che $L\sim Unif(130,135)$
  - La d.d.p. di L è la seguente

$$f(x)=\begin{cases} 0 & L<130 \\ \dfrac{1}{135-130} & 130\leq L\leq 135 \\ 0 & L>135 \end{cases}$$





- La probabilità richiesta è data da:

$$P(132\leq L\leq 134)=\int_{132}^{134} f(x)\,dx=2\cdot\frac{1}{5}=0.40$$

V.c. uniforme: valore atteso e varianza

- Valore atteso di $X \sim Unif(a,b)$: $E[X]=\frac{a+b}{2}$
  - dim.

$$E[X]=\int_{-\infty}^{+\infty} f(x)\,dx=\int_{-\infty}^{a} 0\,dx+\int_{b}^{a}+\frac{1}{b-a}x\,dx+\int_{b}^{+\infty} 0\,dx$$

$$E[X]=\frac{1}{b-a}\int_{a}^{b} x\,dx=\frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b=\frac{1}{b-a}\,\frac{b^2-a^2}{2}=\frac{1}{b-a}\,\frac{(b-a)(b+a)}{2}$$

- Varianza di $X \sim Unif(a,b)$: $Var[X]=\frac{(b-a)^2}{12}$
  - dim.

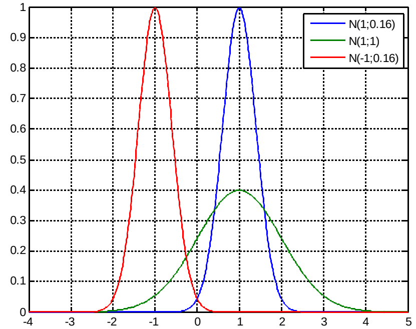
$$Var[X]=\int_{-\infty}^{+\infty} f(x)(x-E[X])^2\,dx=\frac{1}{b-a}\int_{b}^{a}\left(x-\frac{b+a}{2}\right)^2 dx=\frac{1}{b-a}\left[\frac{\left(x-\frac{b+a}{2}\right)^3}{3}\right]_a^b$$

$$Var[X]=\frac{1}{3(b-a)}\left[\left(\frac{2x-b+a}{2}\right)^3\right]_a^b=\frac{1}{3(2^3)(b-a)}\left[(2x-b-a)^3\right]_a^b=\frac{2(b-a)^3}{24(b-a)}$$

V.c. Gaussiana (o normale)

- V.c. continua X le cui modalità appartengono a tutto l'asse reale
- La d.d.p. è nota e dipende da due parametri μ e σ.

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



- Si ha che
  - $E[X] = \mu$
  - $Var[X] = \sigma^2$
  - X è simmetrica rispetto a $E[X]$
- Si dice che: $X \sim N(\mu,\sigma^2)$
- La d.d.p. normale non può (ad oggi) essere integrata in maniera esatta; si può solo calcolare integrali approssimati (come per π).

V.c. Gaussiana (o normale): proprietà

- La trasformazione affine di una gaussiana è ancora una gaussiana
- Esempio: date

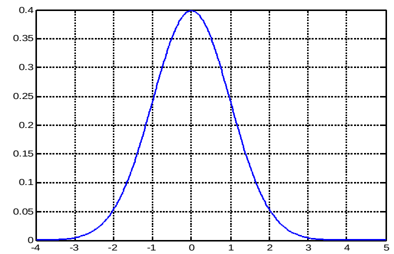
$$X \sim N(\mu,\sigma^2)$$
$$Y = aX + b$$

  - Si ha che
    - $E[X] = \mu$
    - $Var[X] = \sigma^2$
    - $Y \sim N(a\mu+b, a^2\sigma^2)$

- Osservazione: Il terzo punto non è scontato: esistono infinite d.d.p. che condividono lo stesso valore atteso e la stessa varianza.

V.c. Gaussiana (o normale) standardizzata

- Così importante da avere un "nome proprio": Z.
- V.c. Gaussiana in cui si fissano
  - $E[Z] = \mu = 0$
  - $Var[Z] = \sigma^2 = 1$

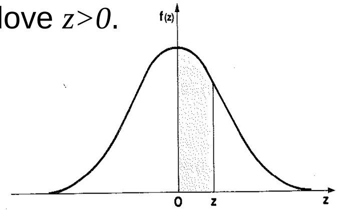


- $Z \sim N(0;1)$.
- Z non ha parametri.
- Gli eventi descritti da una v.c. $X \sim N(\mu;\sigma^2)$ possono essere riferiti a Z. (Processo di standardizzazione)
  - dimostrazione: data $X \sim N(\mu;\sigma^2)$ si consideri $Y = (X-\mu)/\sigma$.
  - $Y \sim N(E[Y];Var[Y])$
  - $E[Y] = (E[X]-\mu)/\sigma = (\mu-\mu)/\sigma = 0.$
  - $Var[Y] = Var[X]/\sigma^2 = \sigma^2/\sigma^2 = 1.$

$\Rightarrow$ $Y \sim Z$

Normale Standardizzata: probabilità

- Gli integrali della d.d.p. di Z son tabulati in tutti i libri di statistica.
- Si conosce $P(0<Z<z)$ dove $z>0$.



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

Ricordando che Z è simmetrica, si ottiene

- $P(0 < Z < z) = P(-z < Z < 0)$.
- $P(Z<z) = 0.5 + P(0 < Z < z)$.
- $P(Z>z) = 1 - P(Z<z) = 0.5 - P(0<Z<z)$.
- $P(Z<-z) = 1 - P(Z<z) = 0.5 - P(0<Z<z)$.
- Dati $z_1 < z_2$; $P(Z < z_1 \cup Z > z_2) = P(Z<z_1) + P(Z>z_2)$.

V.c. normale: calcolo probabilità - I

- Obiettivo: dato x, calcolare $P(X<x)$ dove $X \sim N(\mu;\sigma^2)$.
- Strategia:
  - Standardizzo la modalità x. => $z = (x-\mu)/\sigma$
  - Calcolo (uso le tavole) e trovo $P(Z<z)$.
  - $P(Z<z)$ insiste sullo stesso evento descritto da $P(X<x)$ quindi $P(Z<z) = P(X<x)$.
- Esempio: calcolare $P(X<8)$ dove $X \sim N(10;4)$.
  - Standardizzo la modalità x. => $z = (8-10)/2 = -1$
  - $P(Z < -1) = 0.5 - P(0<Z<1) = 0.5 - 0.3413 = 0.1587$
  - $P(X<8) = P(Z < -1) = 15.87\ \%$

V.c. normale: calcolo probabilità - II

- Obiettivo: dato p, calcolare x: $P(X<x)=p$ dove $X \sim N(\mu;\sigma^2)$.
- Strategia:
  - Calcolo (uso le tavole) e trovo $z : P(Z<z)=p$.
  - $P(Z<z)$ insiste sullo stesso evento descritto da $P(X<x)$ quindi $P(Z<z) = P(X<x)$.
  - Ricavo x da z mediante la trasformazione inversa.
- Esempio: calcolare $x : P(X<x)=0.3$ dove $X \sim N(10;4)$.
  - Poiché $0.3 < 0.5$ si ha che $z < 0$
    $0.3 = 0.5 - P(0<Z<-z)$
    $P(0<Z<-z) = 0.2$ dalle tavole ottengo $-z = 0.525$
  - De-standardizzo la modalità z.

$$Z=\frac{X-E[X]}{\sqrt{Var[X]}} \Leftrightarrow X=\sqrt{Var[X]}Z+E[X] \quad \Rightarrow x=\sqrt{Var[X]}\cdot z+E[X]=8.525$$

Teorema del limite centrale

- Teorema: Date n vv.cc. X_i i.i.d. con $E[X_i] = \mu$ e $Var[X_i] = \sigma^2$, si ha che la v.c.

$$\overline{X}_n=\frac{X_1+X_2+....+X_n}{n}$$

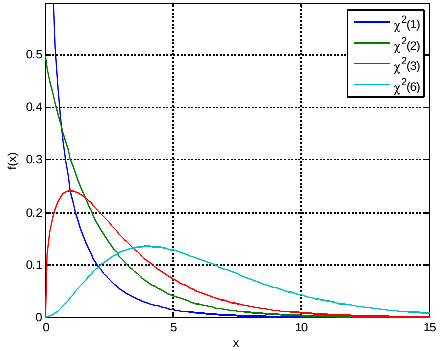
al crescere di n tende a distribuirsi come $N(\mu\ ;\ \sigma^2/n)$.

$$\lim_{n\to\infty} \overline{X}_n \sim N(\mu, \frac{\sigma^2}{n})$$

- *Osservazione:* il teorema non impone il tipo di vv. cc. che vengono mediate, si richiede solo che siano i.i.d. (l'avere lo stesso valore atteso e la stessa varianza è una mera conseguenza).
- *Osservazione:* si può desumere che, non importa che d.d.p. abbia una v.c., se ne prendo abbastanza, la media sarà simile ad una gaussiana.

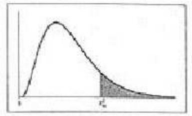
V.c. chi-quadro

- V.c. continua X le cui modalità sono strettamente positive
- D.d.p.
  - nota (ma complessa)
  - nulla per $x \leqslant 0$.
  - 1 parametro $\nu \in \mathbb{N}^+$ (gradi di libertà).
- Si ha che
  - $E[X] = \nu$
  - $Var[X] = 2\nu$
  - X è asimmetrica rispetto a $E[X]$
- Si dice che: $X \sim \chi^2(\nu)$



V.c. chi-quadro: d.d.p.

- d.d.p. Asimmetrica, pertanto non c'è una d.d.p. di riferimento.
- Si usa indicare la d.d.p in funzione
  - dei gradi di libertà $\nu \in \mathbb{N}^+$
  - della probabilità presente nella coda (p-value)



ν	α=0,005	0,01	0,025	0,05	0,100	0,200	0,250	0,500
1	7,88	6,63	5,02	3,84	2,71	1,64	1,32	0,45
2	10,60	9,21	7,38	5,99	4,61	3,22	2,77	1,39
3	12,84	11,34	9,35	7,81	6,25	4,64	4,11	2,37
4	14,86	13,28	11,14	9,49	7,78	5,99	5,39	3,36
5	16,75	15,09	12,83	11,07	9,24	7,29	6,63	4,35
6	18,55	16,81	14,45	12,59	10,64	8,56	7,84	5,35
7	20,28	18,48	16,01	14,07	12,02	9,80	9,04	6,35
8	21,95	20,09	17,53	15,51	13,36	11,03	10,22	7,34
9	23,59	21,67	19,02	16,92	14,68	12,24	11,39	8,34
10	25,19	23,21	20,48	18,31	15,99	13,44	12,55	9,34
11	26,76	24,73	21,92	19,68	17,28	14,63	13,70	10,34
12	28,30	26,22	23,34	21,03	18,55	15,81	14,85	11,34
13	29,82	27,69	24,74	22,36	19,81	16,98	15,98	12,34
14	31,32	29,14	26,12	23,68	21,06	18,15	17,12	13,34
15	32,80	30,58	27,49	25,00	22,31	19,31	18,25	14,34
16	34,27	32,00	28,85	26,30	23,54	20,47	19,37	15,34
17	35,72	33,41	30,19	27,59	24,77	21,61	20,49	16,34
18	37,16	34,81	31,53	28,87	25,99	22,76	21,60	17,34
19	38,58	36,19	32,85	30,14	27,20	23,90	22,72	18,34
20	40,00	37,57	34,17	31,41	28,41	25,04	23,83	19,34

Principio di equivalenza ristretta di Pearson

- Osservazione: Spesso calcolare gli integrali delle d.d.p. è difficile o possibile solo in via numerica.
- Idea: sarebbe bello poter approssimare una d.d.p. ad una di facile integrazione (oppure di integrazione nota).
- Pearson propose il seguente principio:

  Due vv. cc. si dicono ***equivalenti in variabilità*** se i loro indici di forma (simmetria e curtosi) sono uguali.

- Osservazione: quindi se due variabili sono equivalenti in variabilità posso usare la più semplice per il calcolo delle probabilità con un errore trascurabile.

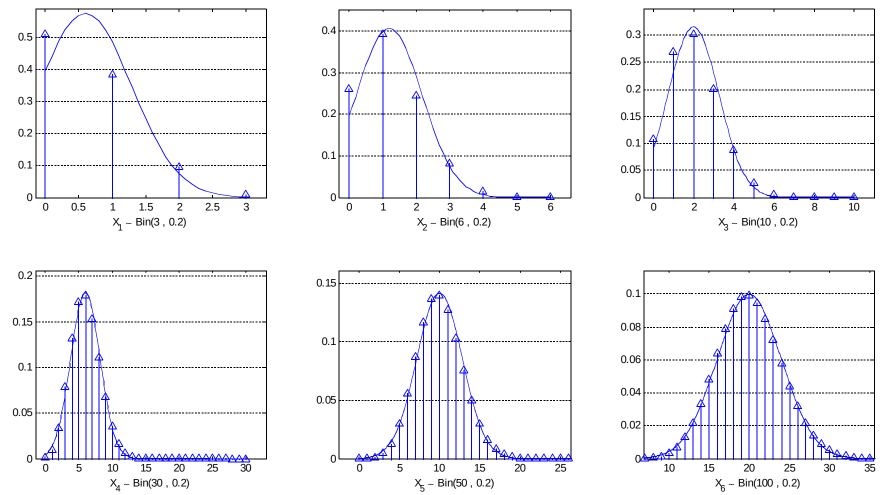
Convergenza in legge: binomiale.

- Si dimostra che una v.c. binomiale al crescere di n "assomiglia" ad una normale con lo stesso valore atteso e varianza.
- Convergenza in legge: il fatto che al divergere di uno o più dei parametri di una d.d.p., essa tenda a sovrapporsi ad un altra
- Distribuzione limite: distribuzione a cui tende una v.c. al crescere di uno (o più) dei suoi parametri.

  Sia $X \sim Bin(n;p)$; se $n \to \infty$, => $X \approx N(np;npq)$

- Osservazione: quanto deve essere grande n perché l'equivalenza valga?
  - Alcuni autori richiedono che $n > 20$.
  - Alcuni autori richiedono che $np > 5$ e $nq > 5$.

Convergenza in legge: binomiale - Esempi.



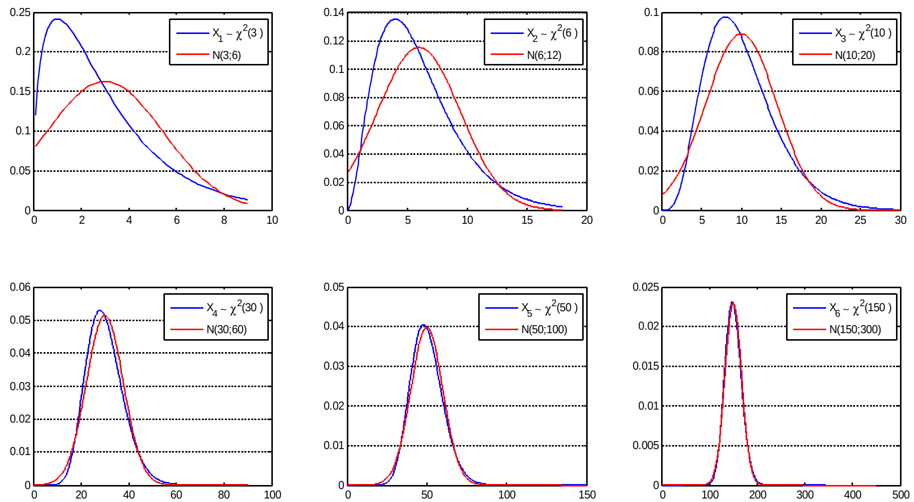
Convergenza in legge: chi quadro.

- La v.c. chi quadro converge in legge ad una normale con lo stesso valore atteso e varianza.

  Sia $X \sim \chi^2(\nu)$; se $\nu \to \infty$, => $X \approx N(\nu;\ 2\nu)$

- Osservazione: quanto deve essere grande ν perché l'equivalenza valga?
  - Alcuni autori richiedono che $\nu > 30$.
  - Alcuni autori richiedono che $\nu > 50$.

Convergenza in legge: chi quadro - Esempi.



Ricapitolando - I

- Densità di probabilità $f(x)$ di una v.c. X: $P(a \le X \le b) = \int_a^b f(x)\,dx$
- Valore atteso: $E[X] = \int_{-\infty}^{+\infty} f(x_i)\,x\,dx$
- Varianza: $Var[X] = \int_{-\infty}^{+\infty} f(x)(x - E[X])^2\,dx$
- Variabili affini: $Y = aX + b$
  - $E[Y] = aE[X] + b$
  - $Var[Y] = a^2 Var[X]$
- Combinazione lineare di vv. cc. indipendenti $Y = \sum_{i=1}^{K} c_i X_i$
  - $E[Y] = \sum_{i=1}^{K} c_i E[X_i]$
  - $Var[Y] = \sum_{i=1}^{K} c_i^2 Var[X_i]$
- Teorema di Bienaymé - Čebičev

$$P(|X - E[X]| > \varepsilon) \le \frac{Var[X]}{\varepsilon^2}$$

Ricapitolando - II

- Uniforme $X \sim Unif(a;b)$.
  - d.d.p. Simmetrica. - .
  - $E[X] = (a+b)/2$; $Var[X] = (b-a)/12$.
- Normale $X \sim N(\mu;\sigma^2)$.
  - d.d.p. Simmetrica.
  - $E[X] = \mu$; $Var[X] = \sigma^2$.
- Normale standardizzata $X \sim N(0;1)$.
- Standardizzazione di $X \sim N(\mu;\sigma^2) \Rightarrow Z = (X - \mu)/\sigma$.
- Chi quadrato. $X \sim \chi^2(\nu)$
  - d.d.p. assimmetrica.
  - $E[X] = \nu$.; $Var[X] = 2\nu$.

Ricapitolando - III

- Teorema del limite centrale:
  date n vv.cc. X_i i.i.d. => $\lim_{n\to\infty} \frac{\sum_{i=1}^{n} X_i}{n} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$
- Equivalenti in variabilità: equivalenza indici di forma.
- Convergenza in legge:
  - $X \sim Bin(n;p)$; se $n \to \infty$, => $X \approx N(np;npq)$
  - $X \sim \chi^2(\nu)$;se $\nu \to \infty$, => $X \approx N(\nu; 2\nu)$