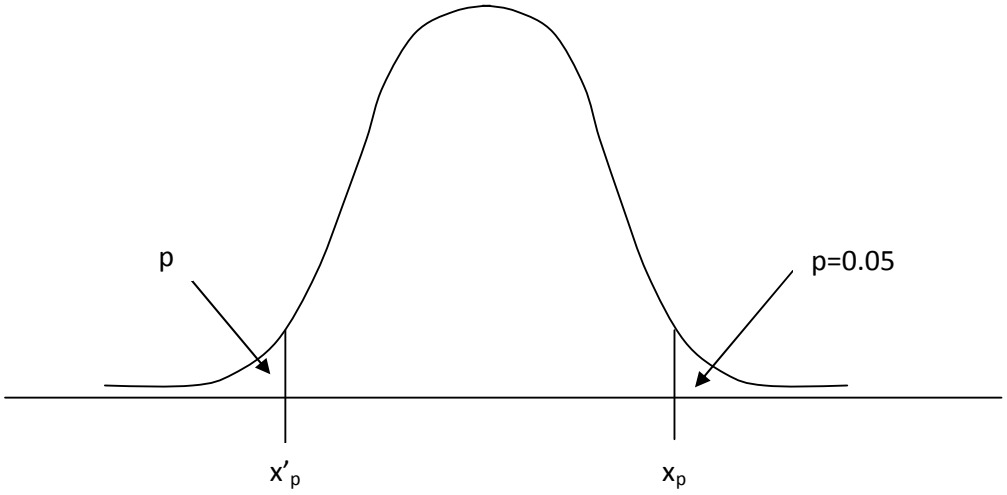
Lezione n. 2 (a cura di Chiara Rossi)

QUANTILE

Data una variabile casuale X, si definisce

Quantile superiore x_p: $X \rightarrow P(X \geq x_p) = p$

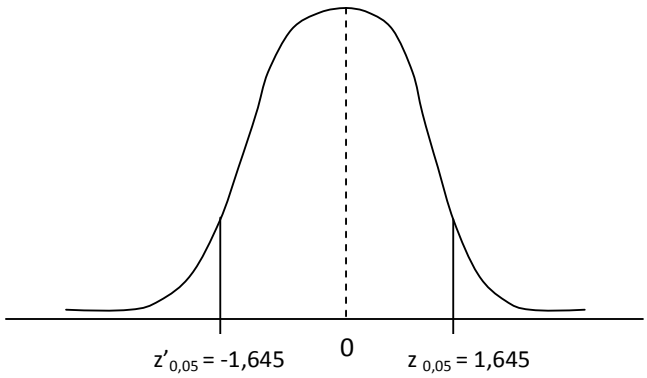
Quantile inferiore x'_p: $X \rightarrow P(X \leq x'_p) = p$



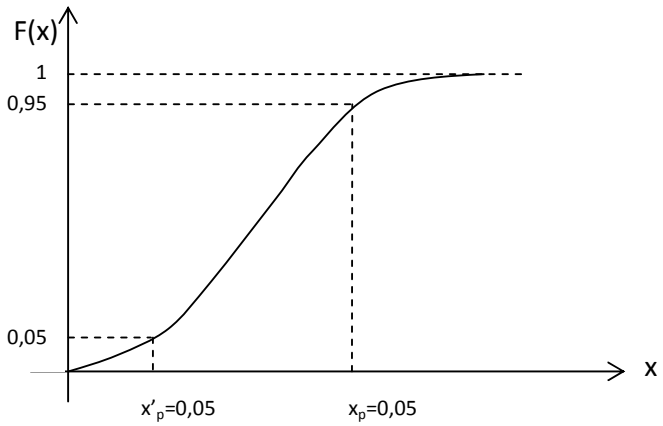
Graficamente, la probabilità p è data dall'area sottostante la curva:

a destra di x_p per il quantile superiore;
a sinistra di x'_p per il quantile inferiore.

Nella distribuzione normale $X \sim N(0,1)$, simmetrica intorno allo 0, il quantile inferiore si trova cambiando di segno il quantile superiore, cioè i due quantili sono simmetrici:



Il quantile si ricava dalla funzione di ripartizione F(x):



Dato che F(x) = P ($X \leq x$), possiamo ricavare direttamente i quantili inferiori; per ottenere i quantili superiori, invece, o li calcoliamo invertendo i quantili inferiori oppure attraverso il complemento ad 1.

FUNZIONE DI DENSITÀ MARGINALE o *DISTRIBUZIONE MARGINALE*

Data una v. c. doppia (X,Y), con *distribuzione marginale* di X intendiamo la distribuzione di X a prescindere da Y.

(X,Y) $\rightarrow$ f(x,y) funzione di densità congiunta

X $\rightarrow$ funzione marginale di X = $f_X(x) = \underbrace{\int f(x,y)\, dy}$

somma rispetto a tutti i valori di y

Y $\rightarrow$ funzione marginale di Y = $f_Y(y) = \int f(x,y)\, dx$

ES: estraggo a caso X = altezza e Y = peso, se dico distribuzione marginale di X intendo SOLO la distribuzione dell'altezza, non mi interessa la distribuzione del peso (e viceversa).

DISTRIBUZIONE CONDIZIONATA

Se abbiamo 2 eventi A e B, quando scriviamo P(A | B) stiamo calcolando la *probabilità condizionata:*

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

Nell'esempio sopra considerato Y | X = "distribuzione del peso data l'altezza".

OSS: con la distribuzione condizionata noi consideriamo solo determinati soggetti (ad es. fissiamo l'altezza a 1,85); mentre con la distribuzione marginale si fa riferimento a tutti i valori di Y.

Nel caso di *variabili doppie continue* esprimiamo il concetto di distribuzione condizionata tramite la funzione di densità:

$$f(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{\text{densità congiunta}}{\text{densità marginale}};$$

le cui caratteristiche sono:

1. $f(y|x) \geq 0$
2. $\int_y f(y|x)\, dy = 1$
3. $P(a < Y < b|X = x) = \int_a^b f(y(x)dy$

Considerando l'esempio sopra, la distribuzione condizionata mi permette di calcolare la probabilità che il peso y di un soggetto sia compreso all'interno di un intervallo [a,b] data l'altezza x.

ES: X = altezza; Y = peso

1. $Y|X = 170 \sim N(65,30)$ utilizzo distribuzione condizionata
   "la distribuzione di Y, dato X = 170, segue una distribuzione normale con media 65 kg e varianza 30"
2. $Y|X = 180 \sim N(75,50)$ utilizzo distribuzione condizionata
3. $Y \sim N(70,40)$ utilizzo distribuzione marginale
4. $P(77 \leq Y \leq 80)$ utilizzo distribuzione marginale
5. $P(77 \leq Y \leq 80|X = 180)$ utilizzo distribuzione condizionata.

INDIPENDENZA

Date due variabili casuali doppie X e Y, *Y è indipendente da X se il valore assunto da X non ha influenza sul valore assunto da Y.*

Si ha indipendenza quando la funzione di densità congiunta è il prodotto delle funzioni marginali:

$$f(x,y) = f_X(x)\, f_Y(y)$$

Se X e Y sono indipendenti allora la distribuzione condizionata coincide con la distribuzione marginale:

$$f(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{f_X(x) f_Y(y)}{f_X(x)} = f_Y(y)$$

Il livello di interdipendenza si misura attraverso l' INDICE DI CORRELAZIONE :

$$\rho = \frac{\mathbf{Cov}\,(\mathbf{X},\mathbf{Y})}{\boldsymbol{\sigma}_{\mathbf{X}}\,\boldsymbol{\sigma}_{\mathbf{Y}}}$$

dove $\sigma_X = \sqrt{\int (x-\mu_x)^2 f_X(x)\,dx}$ e $\sigma_Y = \sqrt{\int (y-\mu_y)^2 f_Y(y)\,dy}$ sono le deviazioni standard, rispettivamente, di X e Y.

$$Cov\,(X,Y) = \iint \underbrace{(x-\mu_X)(y-\mu_Y)}f(x,y)\,dx\;dy$$

differenze tra ogni possibile valore di X e la sua media e tra ogni possibile valore di Y e la sua media

Caratteristiche di ρ :

a. $-1 \leq \rho \leq 1$

se ρ = -1 ⟶ X e Y sono perfettamente correlate negativamente

se ρ = 1 ⟶ X e Y sono perfettamente correlate positivamente

b. $\rho = 0$ ⟶ X e Y sono incorrelate

È importante osservare come il concetto di correlazione sia legato a quello di dipendenza, infatti:

X e Y indipendenti ⇒ X e Y sono incorrelate

X e Y *correlate* ⇒ X e Y sono *dipendenti*

X e Y *incorrelate non* è detto che siano *indipendenti*

ES: ρ = 0,8 ≠ 0 ⇒ c'è correlazione e quindi dipendenza e dato che 0,8 è abbastanza vicino ad 1 la dipendenza si può considerare forte.

VARIABILI CASUALI MULTIPLE

Una v.c. multipla è l'estensione di una v.c. doppia in cui posso osservare n valori $(X_1, X_2, \ldots, X_n)$.

ES: variabili peso, altezza, voto, età ⟶ n = 4 ⟶ 4 dimensioni.

La funzione $f(x_1, x_2, \ldots, x_n)$ è la funzione di densità, con n argomenti quante sono le variabili o dimensioni; ed è:

- non negativa ⟶ $f(x_1, x_2, \ldots, x_n) \geq 0$
- $\int f(x_1, x_2, \ldots, x_n)\,dx = 1$.

COMBINAZIONI LINEARI

Date n variabili casuali $X_1, X_2, \dots, X_n$ si definisce combinazione lineare la seguente variabile casuale:

$$Y = a_1X_1 + a_2X_2 + \dots + a_nX_n$$

dove $a_1, a_2, \dots, a_n$ sono coefficienti fissi.

ES: la media è una combinazione lineare: $E(X) = \frac{1}{n}\sum_i X_i$, $a_i = \frac{1}{n}$, i = 1, 2, … ,n.

Il valore atteso di una combinazione lineare è uguale alla combinazione lineari dei valori attesi:

$$E(Y) = a_1\mu_1 + a_2\mu_2 + \dots + a_n\mu_n$$, dove $\mu_i = E(X_i)$.

La varianza di una combinazione lineare è uguale alla somma delle combinazioni lineari delle varianze e di una costante C:

$$V(Y) = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \dots + a_n^2\sigma_n^2 + C$$, dove $\sigma^2_i = V(X_i)$.

La costante C dipende dalla covarianza, infatti: $C = \sum_i \sum_j a_i a_j Cov(X_i, X_j)$.

Se le variabili casuali sono indipendenti, la Cov = 0 $\Rightarrow$ C sparisce perché è 0.

ES. 1: n=2

$Y = a_1x_1 + a_2x_2$

$E(Y) = a_1\mu_1 + a_2\mu_2$

$V(Y) = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + 2a_1a_2Cov(X_1,X_2)$

ES. 2: n=2, X_1 e X_2 sono indipendenti $\rightarrow$ C = 0

$E(X_1) = 5$ $E(X_2) = 10$

$V(X_1) = 20$ $V(X_2) = 10$

$Y = 2X_1 – X_2$

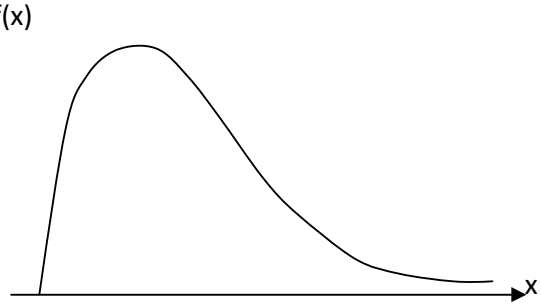
$E(Y) = 2E(X_1) – E(X_2) = 10 – 10 = 0$

$V(Y) = 2^2V(X_1) + V(X_2) = 80 + 10 = 90$

La variabile casuale generata ha media 0 e varianza 90.

DISTRIBUZIONE CHI-QUADRATO:

$X \sim \chi^2(r)$, dove r = gradi di libertà

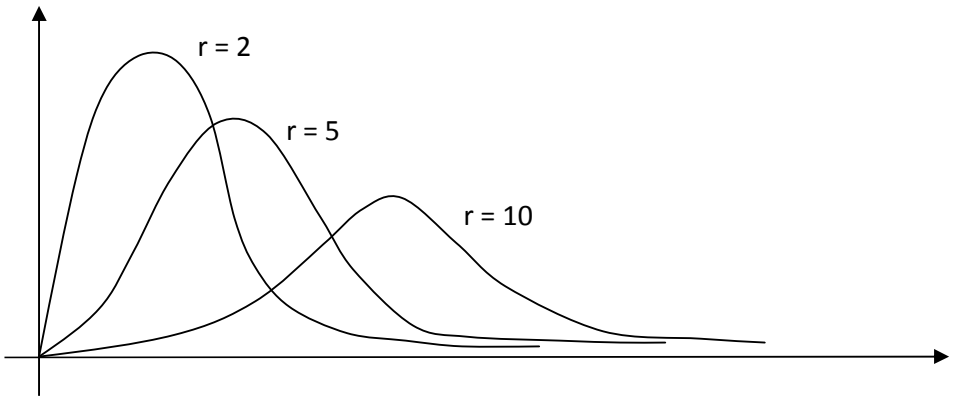


E(X) = r

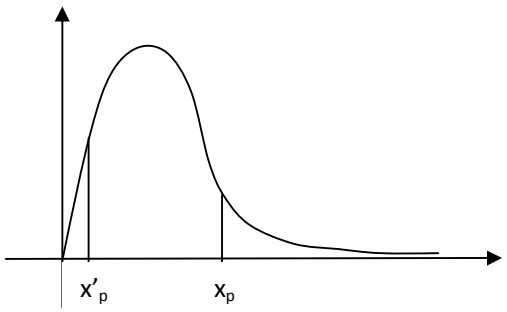
V(X) = 2r

Caratteristiche della distribuzione χ^2:

- è sempre positiva → si applica quando X può assumere soltanto valori positivi (es. reddito);
- è asimmetrica positivamente;
- se r aumenta la curva si sposta verso destra e l'asimmetria decresce, più $r \to \infty$ più la curva tende ad identificarsi con la normale:



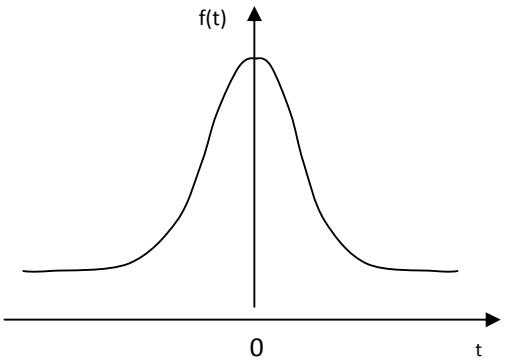
- quantili x_p e x'_p NON sono simmetrici → vanno cercati separatamente



DISTRIBUZIONE T di STUDENT:

Date due variabili casuali $X \sim N(0,1)$ e $Y \sim \chi^2(r)$ indipendenti, allora la trasformazione

$$\frac{X}{\sqrt{\frac{Y}{r}}} = T$$



ha una distribuzione T di Student con r gradi di libertà.

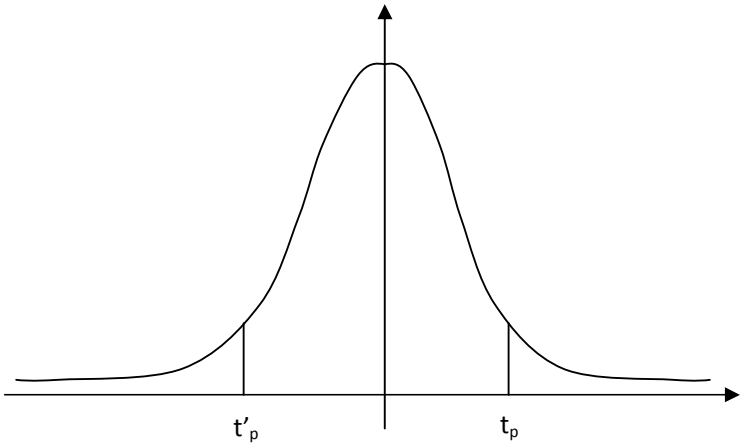
$$\frac{X}{\sqrt{\frac{Y}{r}}} = T \quad \sim t(r)$$

E(T) = 0

V(T) = $\frac{r}{r-2}$, r > 2

Caratteristiche della distribuzione t(r):

- il grafico è simmetrico intorno allo 0;
- se r$\rightarrow\infty$ allora T $\cong$ N (0,1), se r è abbastanza piccolo la forma della t di student non è esattamente uguale a quella della normale;
- quantili t_p e t'_p sono simmetrici e vengono calcolati come i quantili della distribuzione normale:



DISTRIBUZIONE F di Fisher:

Date 2 variabili casuali $X_1 \sim \chi^2(r_1)$ e $X_2 \sim \chi^2(r_2)$ indipendenti, il rapporto

$$\frac{\frac{X_1}{r_1}}{\frac{X_2}{r_2}} = F$$

ha una distribuzione F di Fisher con r_1, r_2 gradi di libertà.

$$\frac{\frac{X_1}{r_1}}{\frac{X_2}{r_2}} = F$$

~ F (r_1,r_2)

La distribuzione F di Fisher è simile alla distribuzione chi-quadrato,

è asimmetrica positivamente

ed i quantili non sono simmetrici e bisogna pertanto calcolarli.