Centro Stampa

N° 4342

SPERIMENTAZIONE ED AFFIDABILITA' DELLE COSTRUZIONI MECCANICHE
TEORIA ESERCIZI 2021-22

DI ARDIZZONI DANIELE

SPERIMENTAZIONE ED AFFIDABILITA' DELLE COSTRUZIONI MECCANICHE

Definizioni preliminari

Teoria della probabilità:
"A branch of mathematics concerned with the analysis of random phenomena" [Encyclopaedia Britannica].
Parte della matematica che, sulla base delle informazioni a disposizione, sviluppa metodi e calcoli per esprimere quantitativamente la tendenza di certi eventi a verificarsi.

Statistica:
Scienza che studia con metodi matematici fondati sul calcolo delle probabilità fenomeni collettivi e di massa [Devoto-Oli].

Popolazione:
Insieme di elementi considerati omogenei rispetto ad una o più caratteristiche.

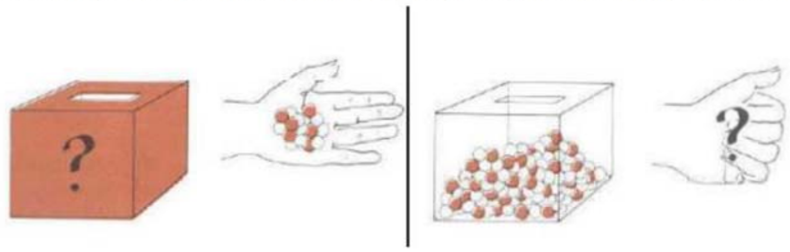
- Studenti in aula (popolazione finita di numerosità ridotta).
- Lotto di produzione: Bulloni nominalmente uguali prodotti in 10 anni (popolazione finita di numerosità elevata).
- Tempi di attesa per un servizio (popolazione infinita).

Elemento:

- Ogni singolo studente in aula.
- Ogni singolo bullone prodotto
- Ogni tempo di attesa realizzabile (t>0).

Poi abbiamo le applicazioni della statistica
Statistica descrittiva:
L'insieme dei metodi che riguardano la raccolta, la presentazione e la definizione di un insieme di dati per descriverne in maniera adeguata le varie caratteristiche.

Statistica inferenziale:
L'insieme dei metodi che permettono la stima di una caratteristica della popolazione e che sono basati soltanto su una sua parte (opportunamente scelta), detta "campione". Si prende un campione dalla popolazione in maniera casuale e indipendente (indipendente vuol dire che quando estraggo un elemento non cambio la probabilità di un altro elemento di essere estratto), facciamo lo studio del campione e a seconda di questo studio cerchiamo di capire qualcosa sulla popolazione.



Statistics:
Given the information
in your hand,
what is the box?

Probability:
Given the information
in the box,
what is in your hand?

Variabile casuale (o aleatoria)

Grandezza scalare, ai suoi valori numerici viene associata una legge di probabilità che esprime una misura della loro tendenza a verificarsi.

Variabile casuale discreta (valori numerabili):

n_a= frequenza assoluta dell'evento a; n = risultati

Frequenza relativa: $f_a = \frac{n_a}{n}$

Probabilità (a posteriori): $p(a) = \lim_{n \to \infty} \frac{n_a}{n}$

Esiste anche la definizione assiomatica basata sui 3 assiomi di Kolmogorov valida anche per variabile casuale continua.

1) $0 \leq p(a) \leq 1$

Evento impossibile Evento certo

I e II assioma di Kolmogorov

2) Eventi mutuamente esclusivi:

$$p(a+b) = p(a \cup b) = p(a) + p(b)$$

probabilità di avere l'evento *a* oppure l'evento *b*

III assioma di Kolmogorov

Esempio, se ci sono 3 sfere di diverso colore in un sacchetto si vuole trovare la probabilità di estrarre una sfera di un dato colore

3) Eventi indipendenti

$$p(ab) = p(a \cap b) = p(a) \cdot p(b)$$

Probabilità che si verifichino entrambi gli eventi

Esempio, se lanciamo una moneta probabilità di avere due teste consecutive

4) Eventi indipendenti ma non mutamente esclusivi

$$p(a+b) = p(a \cup b) = p(a) + p(b) - p(a \cap b)$$

Esempio, due sfere, una rossa e una bianca, due cubi, uno rosso e uno bianco, la probabilità di estrarre un cubo o un oggetto bianco. Oppure la probabilità di estrarre da un mazzo di carte un cuore o un asso

5) Eventi dipendenti (probabilità condizionata)

$$p(a/b) = \frac{p(a \cap b)}{p(b)} \Rightarrow p(ab) = p(a \cap b) = p(b) \cdot p(a/b)$$

Probabilità che si presenti l'evento *a* essendosi verificato l'evento *b*

Esempio, tre palline bianche, due palline nere, due estrazioni senza reintroduzione, la probabilità che la seconda pallina sia bianca se la prima pallina estratta è bianca

b) Prima pallina estratta bianca: p(b)=3/5

a) seconda pallina bianca: p(a/b) =1/2

$$p(ab) = p(a \cap b) = p(b) \cdot p(a/b) = \frac{3}{5} \cdot \frac{1}{2} = \frac{3}{10}$$

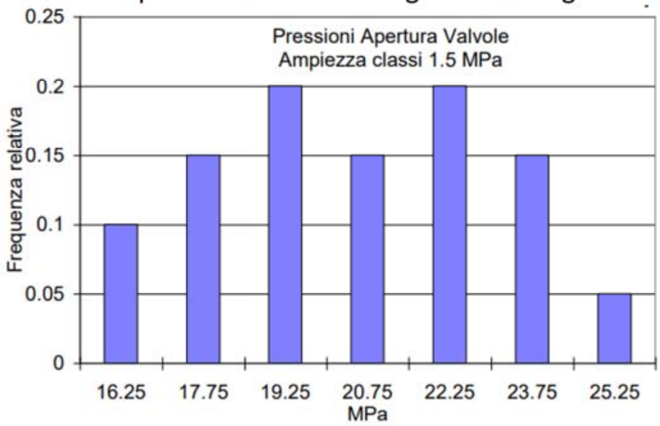
È chiaro che se la numerosità del campione è molto grande la probabilità condizionata tende agli eventi indipendenti. Nel lotto la probabilità di estrarre uno dei 90 numeri è 1/90, la probabilità di estrarre un altro numero è 1/89, ma noi abbiamo la probabilità condizionata e quindi è 1/89 * 89/90 e quindi la probabilità di trovare tutti i numeri è 1/90 ^5.

Variabile casuale continua

Il numero di possibili risultati è infinito. La probabilità di ottenere un preciso risultato è praticamente nulla. Allora per poter ricavare dei risultati utili conviene organizzare i dati in classi definite da intervalli $[x_i - x_{i+1}]$ di ampiezza d_i.

Frequenza relativa alla classe *i:* $f_i = \frac{n_i}{n}$ Densità: $\delta_i = \frac{f_i}{d_i} = \frac{n_i}{n \cdot d_i}$

I dati sono presentati tramite istogrammi o diagrammi a barre



La frequenza relativa alla classe viene presa relativa perché (come esempio) un conto è prendere la probabilità che un componente si rompa in un mese, un conto è in una settimana.

Funzione di densità di probabilità (*fdp*):

$$f(x) = \lim_{\substack{n \to \infty \\ d_i \to 0}} \delta_i = \lim_{\substack{n \to \infty \\ d_i \to 0}} \frac{n_i}{n \cdot d_i} \qquad \int_{-\infty}^{+\infty} f(x)dx = 1$$

La *fdp* **non esprime la probabilità** che la variabile assuma il valore *x*

È la probabilità di avere la variabile che cada all'interno dell'intervallo nell'intorno di x, per fare delle considerazioni sulla probabilità si usa la:

Funzione di probabilità cumulata $F(a) = \int_{-\infty}^{a} f(x)dx = p(x < a)$

Proprietà:

1) $F(-\infty) = 0$ evento impossibile
2) $F(+\infty) = 1$ evento certo
3) $p(a < x < b) = F(b) - F(a)$
4) $p(x > a) = 1 - F(a)$

È la probabilità di avere un valore più piccolo del valore cercato. La probabilità all'estremo inferiore è l'evento impossibile (la probabilità di avere una persona alta 0 cm è nulla), (la probabilità che una persona sia più bassa di 3 metri è certa), (la probabilità di avere l'altezza di una persona fra 179,5 cm e 180,5 cm è quella di avere la probabilità di 180 cm perché sulle carte di identità l'altezza è espressa in cm e la probabilità di essere più bassi è 1- la probabilità di essere superiore a 180 cm).

Parametri di posizione per dei campioni

Media aritmetica: $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{j=1}^{K} f_j x_j}{n}$

Media geometrica $G = \sqrt[n]{\prod x_i}$ Media armonica $H = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}}$

Mediana: Valore centrale (n dispari); Media dei valori centrali (n pari)

Moda: Valore con frequenza maggiore

Parametri di posizione per delle popolazioni

Media o valore atteso
$$\mu = E(x) = \sum x_i p(x_i) \quad \text{var. discrete}$$
$$\mu = E(x) = \int_{-\infty}^{+\infty} x \cdot f(x) \cdot dx \quad \text{var. continue}$$

Mediana (x_{50}) Valore per cui $F(x) = 0.5$

Moda Valore per cui $f(x) = \max$

Percentili Valori che dividono in cento parti uguali i dati

10° percentile = $F(x) = 0.1$

Si usano anche decili, i quartili eccetera.
La media è solitamente chiamata "valore atteso" perché si esprime come la somma dei prodotti dei valori per la probabilità di avere quei valori.

Parametri di variabilità per dei campioni

Si usano perché abbiamo una certa distribuzione dei nostri campioni

Campioni

Scarto quadratico medio (deviazione standard)
$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{\sum x_i^2 - n \cdot \bar{x}^2}{n-1}}$$

Varianza $V = s^2$

Si divide per $n - 1$ perché (si può dimostrare) quando ho una serie di dati, si può fare una considerazione sui dati anche con la numerosità, se ho due punti ci passa una sola retta, quando faccio la media tolgo un "grado di libertà" e per questo si calcola con $n - 1$. La varianza è lo scarto quadratico medio al quadrato.

Parametri di variabilità per delle popolazioni

Varianza
$$\sigma^2 = E((x-\mu)^2) = \frac{\sum_{1}^{\infty}(x_i-\mu)^2}{n}$$

Scarto quadratico medio (deviazione standard - popolazione)
$$\sigma = \sqrt{\sigma^2}$$

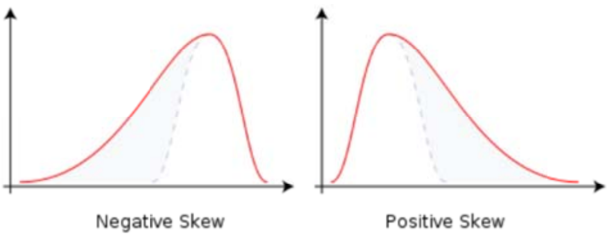
per una popolazione invece varianza e deviazione standard si dividono per n e non per $n-1$.

Asimmetria

In statistica l'asimmetria si misura con l'indice di asimmetria a:

$$a = \frac{\sum(x_i-\overline{x})^3}{n \cdot s^3} \quad \text{(campione)} \qquad a = E\left[\left(\frac{(x_i-\mu)}{\sigma}\right)^3\right] \text{(popolazione)}$$

$a = 0 \Rightarrow$ distribuzione simmetrica
$a < 0 \Rightarrow$ i dati minori della media sono più lontani
$a > 0 \Rightarrow$ i dati maggiori della media sono più lontani



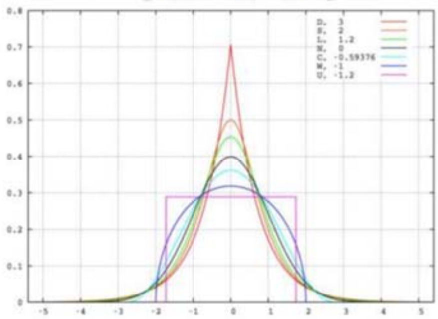
L'asimmetria si adimensionalizza con lo scarto tipo elevato al cubo.

Curtosi

La curtosi è l'allontanamento dalla normalità distributiva, cioè dalla forma di distribuzione gaussiana.

$$k = \frac{\sum(x_i-\overline{x})^4}{n \cdot s^4} \text{(campione)} \qquad k = E\left[\left(\frac{(x_i-\mu)}{\sigma}\right)^4\right] \text{(popolazione)}$$

$k \approx 3 \Rightarrow$ mesocurtosi (come distribuzione normale)
$k < 3 \Rightarrow$ leptocurtosi (dispersione minore della normale)
$k > 3 \Rightarrow$ platicurtosi (dispersione maggiore della normale)



$\longleftarrow$ k-3

La campana di distribuzione di Gauss ha una curtosi $k = 3$, gli altri dati della curtosi portano la campana ad assumere forme più o meno piatte.

Se abbiamo due variabili (x e y) si possono inoltre definire

<u>Covarianza</u>

$$\mathrm{cov}(x,y) = E\left[(x-\mu_x)(y-\mu_y)\right]$$
$$\mathrm{cov}(x,y) = \sum\left(x-\bar{x}\right)\left(y-\bar{y}\right)/n \;= s_{xy}$$

proprietà

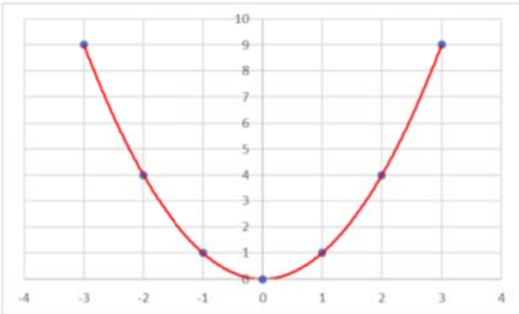
$$\mathrm{cov}(x,a) = 0 \qquad \mathrm{cov}(x,y) = \mathrm{cov}(y,x)$$
$$\mathrm{cov}(x,x) = \sigma_x^2 \qquad \mathrm{cov}(ax,by) = ab\,\mathrm{cov}(x,y)$$
$$\mathrm{cov}(x+a, y+b) = \mathrm{cov}(x,y)$$

La covarianza è "come varia una cosa rispetto a un'altra". Se c'è una covarianza molto alta non è detto che i due siano collegati, correlazione non implica causalità.

<u>Correlazione</u>

Coefficiente di Pearson: misura della dipendenza lineare fra le due variabili

$$\rho = \frac{\mathrm{cov}(x,y)}{\sigma_x \sigma_y} \qquad -1 \le \rho \le 1$$



Se x and y sono indipendenti, la covarianza e il coefficiente di correlazione sono zero, ma non è vero il contrario perché il coefficiente di Pearson riguarda solo la dipendenza lineare fra x e y variabili, quindi le dipendenze paraboliche non vengono considerate.

<u>Variabile casuale standardizzata</u>

Il problema è valutare se ci sono delle leggi di distribuzione teoriche, prima di guardare i modelli, iniziamo a definire una variabile casuale standardizzata. Medie, varianze, dipendono dalle unità di grandezza, noi vogliamo adimensionalizzare e centrare, allora introduciamo la variabile casuale standardizzata, i simboli sono da notare, di solito si indica con la z ma noi per comodità utilizzeremo sempre la y.

$$y_i = \frac{x_i - \bar{x}}{s} = z_i \qquad \bar{y} = 0 \quad s_y = 1$$

Il centrare e l'adimensionalizzazione ci permette di avere una media nulla e una deviazione standard uguale a 1.

Nel cambiamento di variabile non devono cambiare le probabilità

$$f(x)dx = f(y)dy \qquad dy = \frac{dx}{s} \Rightarrow f(x) = \frac{f(y)}{s}$$

$$\Downarrow$$

$$F(x_h) = F(y_h)$$

Che io calcoli la funzione con il valore diretto o quello standardizzato abbiamo lo stesso risultato.

Modelli teorici variabili casuali

Distribuzione normale (o Gaussiana) Campo di definizione $-\infty; +\infty$.
Parametri 2

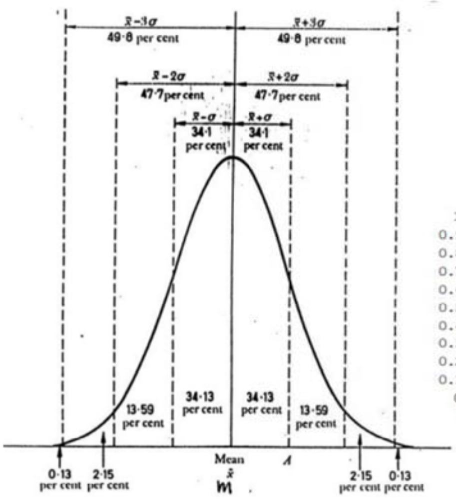
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \qquad f(y) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{y^2}{2}\right]$$

$$F(a) = \int_{-\infty}^{a} f(x)dx$$ Non integrabile analiticamente

- $f(x)$ simmetrica
- flessi per $x = \mu \pm \sigma$
- media moda e mediana coincidono: $F(\mu)=0.5$

$$F(\mu - b) = 1 - F(\mu + b)$$
$$F(-y) = 1 - F(y) \quad F(y) = 1 - F(-y)$$

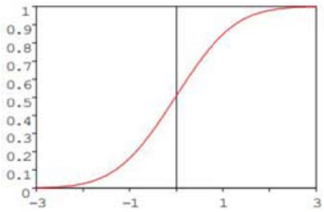
Si chiama *normale* perché molte popolazioni naturali seguono questo andamento a campana. La caratteristica scomoda di questa distribuzione è che non è integrabile analiticamente e quindi bisogna utilizzare delle tabelle.



Tabelle

$\mu - \sigma < x < \mu + \sigma$ 68.26%

$\mu - 2\sigma < x < \mu + 2\sigma$ 95.4%



Distribuzione lognormale

È la distribuzione normali dei logaritmi

NB: $t \neq t_{50}$

$$x_l = \ln(t) \qquad dx_l = \frac{dt}{t} \qquad f(x_l) = \frac{1}{\sigma_l\sqrt{2\pi}} \exp\left[-\frac{(x_l-\mu_l)^2}{2\sigma_l^2}\right]$$
$$0 \le t \le \infty \qquad -\infty \le x_l \le +\infty$$

$$F(t) = F(x_l) = F(y) \Rightarrow f(t) = \frac{1}{\sigma_l t\sqrt{2\pi}} \exp\left[-\frac{(\ln(t)-\mu_l)^2}{2\sigma_l^2}\right]$$

Nel caso in cui si considerino i logaritmi in base 10 si ha:

$$f(t) = \frac{1}{\ln(10)\sigma_l t\sqrt{2\pi}} \exp\left[-\frac{(\log(t)-\mu_l)^2}{2\sigma_l^2}\right]$$

$\log_a c \cdot \log_c b = \log_a b$
$\ln(10) \cdot \log(b) = \ln(b)$

Se fossero logaritmi in base decimale basta considerare le regole di trasformazione della base dei logaritmi.

Alcune formule utili

Se io conosco la media e la varianza dei logaritmi e io voglio passare a quelli normali, si può dimostrare che ci sono delle formule di trasformazioni

$$\mu_t = \exp\left(\mu_l + \frac{\sigma_l^2}{2}\right)$$

$$\sigma_t^2 = \exp(2\mu_l + \sigma_l^2)(\exp(\sigma_l^2) - 1)$$

Nel caso si abbiano i valori in log decimali:

$$\mu_l = \ln(10) \cdot \mu_{l10}$$
$$\sigma_l = \ln(10) \cdot \sigma_{l10}$$

$$\sigma_l^2 = \ln\left(\frac{\sigma_t^2}{\mu_t^2} + 1\right)$$

$$\mu_l = \ln(\mu_t) - \frac{\sigma_l^2}{2}$$

Volendo i valori in log decimali:

$$\mu_{l10} = \frac{\mu_l}{\ln(10)}$$

$$\sigma_{l10} = \frac{\sigma_l}{\ln(10)}$$

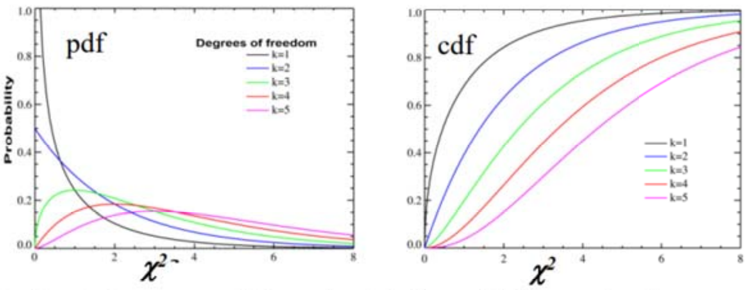
Distribuzione chi-quadro χ^2

È una distribuzione per capire se una certa variabile in una distribuzione normale ha dei legami con un'altra variabile.

Se $y_1, \ldots, y_k$ sono variabili standardizzate distribuite normalmente ed indipendenti:

$$Q = \sum_{i=1}^{k} y_i^2$$

È distribuita secondo la distribuzione quadro χ^2 con k gradi di libertà $Q \approx \chi^2(k)$



Se faccio la somma dei quadrati delle variabili standardizzate ottengo un valore Q, questo valore Q segue una distribuzione χ^2, questa distribuzione dipende dal grado di libertà.

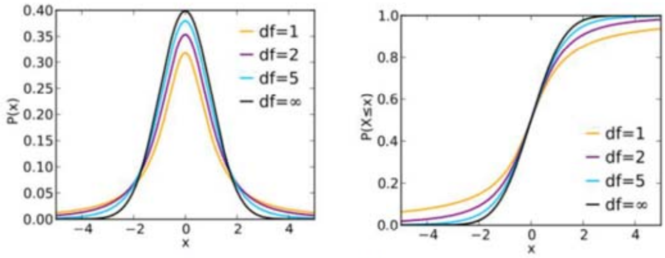
Il numero di gradi di libertà è il numero di valori di una statistica che è libero di variare. Se abbiamo un solo dato, Q vale come quel valore, all'aumentare degli elementi del campione, cambia la distribuzione.

(la cdf è la cumulative distribution function ed è "l'accumulo" della probabilità che un evento sia quello atteso, e infatti tende verso a 1).

Distribuzione t di Student

La distribuzione t di Student è la pdf (pdf probability dencity function e fdp funzione della probabilità sono la stessa cosa) del rapporto:

$$\frac{y}{\sqrt{Q/\nu}}$$



Per $n \geq 30$ è indistinguibile dalla distr. normale

y è distribuita normalmente con media 0 e varianza 1. Q è distribuita secondo la distribuzione chi-quadro con ν gdl. y e Q sono indipendenti.

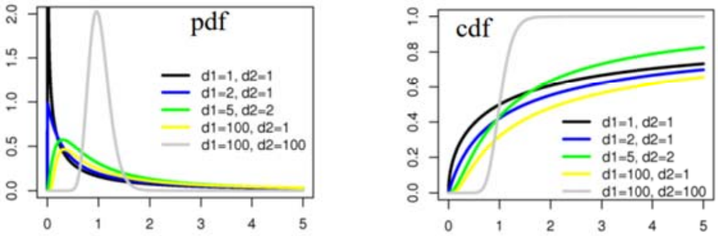
"Student's t-distribution arises in the problem of estimating the mean of a normally distributed population when the sample size is small (see later).
Student was the pen name of William Sealy Gosset, that joined the Dublin brewery of Arthur Guinness & Son after a graduation in Oxford. Guinness was a progressive agro-chemical business and Gosset would apply his statistical knowledge both in the brewery and on the farm—to the selection of the best yielding varieties of barley. Another researcher at Guinness had previously published a paper containing trade secrets of the Guinness brewery. To prevent further disclosure of confidential information, Guinness prohibited its employees from publishing any papers regardless of the contained information. This meant that Gosset was unable to publish his works under his own name."

Distribuzione F (di Snedecor o di Fisher-Snedecor)

Questa distribuzione F è fondamentale perché riguarda le varianze di due popolazioni, capire se sono uguali sulla base delle varianze campionarie (di due campioni). Abbiamo introdotto la χ^2 perché la distribuzione F è il rapporto fra le due χ^2 fratto il numero di gradi di libertà. A seconda del numero di gradi di libertà cambia la funzione. Ci sono due gradi di libertà e una probabilità da calcolare.

Se χ^2_x e χ^2_y sono due variabili casuali indipendenti che seguono una distribuzione χ^2 con ν_x e ν_y gdl, la variabile:

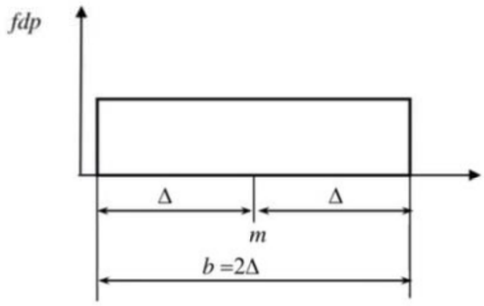
$$\frac{\chi^2_x/\nu_x}{\chi^2_y/\nu_y} \quad \text{Segue una distribuzione } F \text{ con } \nu_x, \nu_y \text{ gdl}$$



$$se \quad X \approx F(\nu_1, \nu_2) \Rightarrow \frac{1}{X} \approx F(\nu_2, \nu_1)$$

Se un valore è la funzione F con due gradi di libertà, se scambio di valori dei dati di libertà al posto di X abbiamo $\frac{1}{X}$ una specie di simmetria.

Distribuzione uniforme



$$\sigma^2 = \frac{b^2}{12} = \frac{\Delta^2}{3} \quad \sigma = \frac{b}{\sqrt{12}} = \frac{\Delta}{\sqrt{3}}$$

La media vale il valore centrale, la varianza è la distanza fra i due estremi diviso $\sqrt{12}$. Questo si può utilizzare nello scarto di lavorazione quando abbiamo delle tolleranze di misura.

Distribuzione di Weibull

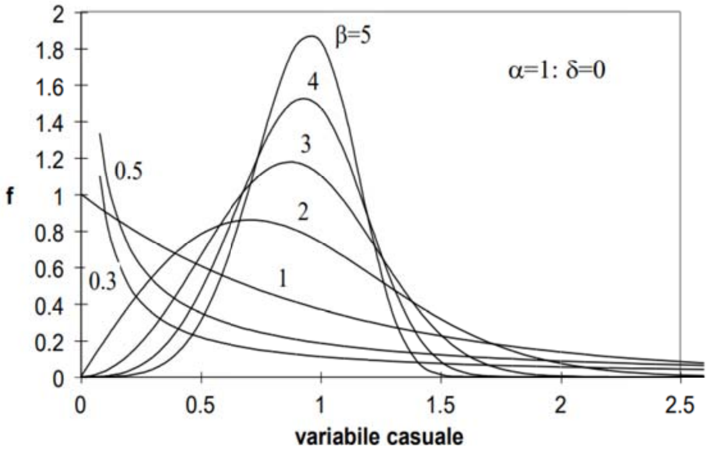
Anche Weibull lavorava per una azienda (SKF) e doveva delle stime sulla affidabilità. Formulò una teoria sulla distribuzione di resistenza dei materiali. Trovò una distribuzione che non dipendesse da medie e varianze.

$$F(t) = 1 - \exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right] \qquad f(t) = \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1}\exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right]$$

$$F(t) = 1 - \exp\left[-\left(\frac{t-\delta}{\alpha}\right)^{\beta}\right] \qquad f(t) = \frac{\beta}{\alpha}\left(\frac{t-\delta}{\alpha}\right)^{\beta-1}\exp\left[-\left(\frac{t-\delta}{\alpha}\right)^{\beta}\right]$$

t = variabile
α = fattore di scala o vita caratteristica
β = fattore di forma
δ = valore minimo garantito ($t' = t - \delta$); $F(\delta) = 0$

$$F(t' = \alpha) = 1 - \exp(-1) = 0.632 \quad \forall\beta$$

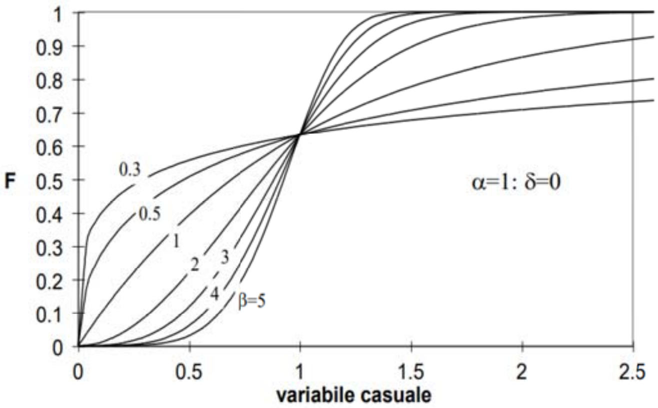


La variabile t va da 0 a ∞, è una durata in pratica, i valori di α *e* β si calcolano diversamente da media e varianza. δ è un altro parametro (infatti si chiama anche "distribuzione a tre parametri") ed è il valore di resistenza minimo del materiale (nella ricerca di Weibull).

In corrispondenza della vita caratteristica i componenti che si rompono sono il 63,2%.

Al variare di β cambia la forma, a $\beta = 3$ abbiamo più o meno la campana di Gauss.

La funzione di probabilità cumulata



Calcolo dei percentili

Nella distribuzione di Weibull, se io voglio conoscere quanto vale il valore della variabile t, basta calcolare, la funzione di probabilità cumulata conoscendo i tre parametri mettendo dentro t_P (che è l'incognita) e la $F(t_P)$ è il valore che abbiamo perché lo imponiamo noi (per esempio il decimo percentile =10%).

$$F(t_p) = 1 - \exp\left[-\left(\frac{t_p - \delta}{\alpha} \right)^{\beta} \right]$$

$$1 - F(t_p) = \exp\left[-\left(\frac{t_p - \delta}{\alpha} \right)^{\beta} \right]$$

$$\ln(1 - F(t_p)) = -\left(\frac{t_p - \delta}{\alpha} \right)^{\beta}$$

$$\frac{t_p - \delta}{\alpha} = \left(\ln \frac{1}{1 - F(t_p)} \right)^{\frac{1}{\beta}}$$

$$t_p = \delta + \alpha \left(\ln \frac{1}{1 - F(t_p)} \right)^{\frac{1}{\beta}} = t_p = \delta + \alpha\left(-\ln(1 - F(t_p))\right)^{\frac{1}{\beta}}$$

<u>Distribuzione esponenziale</u>

È un caso particolare della Weibull a due parametri con $\beta = 1$

$$F(t) = 1 - \exp\left[-\frac{t}{\alpha} \right] \qquad f(t) = \frac{1}{\alpha} \exp\left[-\frac{t}{\alpha} \right]$$

NB: $$E(t) = \int_{-\infty}^{\infty} t \cdot f(t) \cdot dt = \int_{-\infty}^{\infty} t \cdot \frac{1}{\alpha} \exp\left[-\frac{t}{\alpha} \right] \cdot dt = \alpha$$

<u>Distribuzioni dei valori estremi</u>

Distribuzioni seguite dai valori minimi (massimi) di campioni - di numerosità tendente ad infinito - ottenuti da popolazioni che seguono distribuzioni Normali, log-normali o di Weibull.
A noi potrebbe interessare come si distribuiscono i valori più piccoli (o più grandi, come ad esempio la resistenza minima o massima di una molla). Se abbiamo dei problemi di fatica e dobbiamo capire qual è il minimo del limite di fatica.

<u>Distribuzione di Bernoulli (BINOMIALE) - variabili discrete</u>

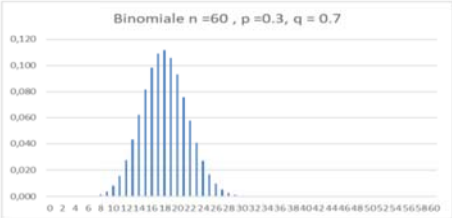
Tipica distribuzione dei controlli di qualità (ma anche dei giochi...) Supponiamo di estrarre n pezzi da una popolazione in cui il p è la proporzione di pezzi difettosi e q = 1-p la proporzione di pezzi buoni. La probabilità di estrarre r pezzi difettosi è:

$$P(r) = \binom{n}{r} p^r q^{n-r} \quad \text{dove} \quad \binom{n}{r} = \frac{n!}{r!\,(n-r)!}$$

$$n! = 1 \cdot 2 \cdot 3 \cdot \ldots\ldots \cdot n$$
$$0! = 1$$

Se il numero di estrazioni n è sufficientemente grande la distribuzione Binomiale può essere approssimata da una distribuzione Normale con

$$\mu = np \qquad \sigma^2 = npq$$



p	0,1		r	p(r)	P(r)
q	0,9		0	34,87%	34,87%
n	10		1	38,74%	73,61%
			2	19,37%	92,98%
			3	5,74%	98,72%
			4	1,12%	99,84%
			5	0,15%	99,99%
			6	0,01%	100,00%
			7	0,00%	100,00%
			8	0,00%	100,00%
			9	0,00%	100,00%
			10	0,00%	100,00%
			Totale	1,000	

Regole empiriche: l'approssimazione è accettabile se:

$$np > 5 \ e \ nq > 5$$

Oppure

$$npq > 9$$

Questo ci dice che nel caso del controllo qualità per trovare la probabilità che un pezzo esca rotto non basta fare il rapporto fra pezzi buoni e rotti, perché anche la probabilità di rottura si distribuisce con una certa probabilità

Distribuzione di Poisson (variabili discrete)

Descrive la probabilità del numero di eventi in un continuo. Esempio: foratura pneumatici.

y = n° medio di eventi in un determinato tempo $\mu = y \quad \sigma = \sqrt{y}$

$$P(r) = \frac{e^{-y} y^r}{r!}$$

Esempio: prove di detonazione. 10 motori monocilindrici. Durata 30'.
Unità di misura tempo = 1 minuto

Det/minuto	Nv	nd x nv		r	p(r)	fr. Teorica
0	150	0		0	48,2%	144,57
1	100	100		1	35,2%	105,54
2	36	72		2	12,8%	38,52
3	10	30		3	3,1%	9,37
4	3	12		4	0,6%	1,71
5	1	5		5	0,1%	0,25
	minuti	tot				
	300	219				
	y=	0,730				

È comodo perché ha una distribuzione continua, e ci permette di calcolare la probabilità ad andare a infinito, quindi la vera distribuzione.

Algebra normale

2 variabili, x e y

$$\mu_{ax\pm by} = a\mu_x \pm b\mu_y \quad \sigma^2_{ax\pm by} = a^2\sigma_x^2 + b^2\sigma_y^2$$

Se x e y distribuite normalmente ($ax\pm by$) è ancora normale

$$\mu_{xy} = \mu_x\mu_y \quad \sigma^2_{xy} = \mu_x^2\sigma_y^2 + \mu_x^2\sigma_y^2 + \sigma_y^2 \cdot \sigma_x^2$$

$$\mu_{x/y} \cong \frac{\mu_x}{\mu_y} \quad \sigma^2_{x/y} \cong \left(\frac{\mu_x}{\mu_y}\right)^2 \left(\frac{\sigma_x^2}{\mu_x^2} + \frac{\sigma_y^2}{\mu_y^2}\right)$$

Queste formule valgono per le distribuzioni normali, se abbiamo due variabili x e y, la media è la somma di queste due distribuzioni, la varianza è sempre la somma delle varianze (propagazione degli errori).
Quando la media è un prodotto o un rapporto fra le medie la varianza sarà ottenuta algebricamente come nelle formule viste sopra.

Vediamo ora la rappresentazione di una Gaussiana su Excel

G13 fx =DISTRIB.NORM.ST.N($A13+G$2;FALSO)

	A	B	C	D	E	F	G	H	I	J	K
1		Gaussiana densità di probabilità			DISTRIB.NORM.ST.N(y; FALSO)						
2	y	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
3	0,0	3,9894E-01	3,9892E-01	3,9886E-01	3,9876E-01	3,9862E-01	3,9844E-01	3,9822E-01	3,9797E-01	3,9767E-01	3,9733E-01
4	0,1	3,9695E-01	3,9654E-01	3,9608E-01	3,9559E-01	3,9505E-01	3,9448E-01	3,9387E-01	3,9322E-01	3,9253E-01	3,9181E-01
5	0,2	3,9104E-01	3,9024E-01	3,8940E-01	3,8853E-01	3,8762E-01	3,8667E-01	3,8568E-01	3,8466E-01	3,8361E-01	3,8251E-01
6	0,3	3,8139E-01	3,8023E-01	3,7903E-01	3,7780E-01	3,7654E-01	3,7524E-01	3,7391E-01	3,7255E-01	3,7115E-01	3,6973E-01
7	0,4	3,6827E-01	3,6678E-01	3,6526E-01	3,6371E-01	3,6213E-01	3,6053E-01	3,5889E-01	3,5723E-01	3,5553E-01	3,5381E-01
8	0,5	3,5207E-01	3,5029E-01	3,4849E-01	3,4667E-01	3,4482E-01	3,4294E-01	3,4105E-01	3,3912E-01	3,3718E-01	3,3521E-01
9	0,6	3,3322E-01	3,3121E-01	3,2918E-01	3,2713E-01	3,2506E-01	3,2297E-01	3,2086E-01	3,1874E-01	3,1659E-01	3,1443E-01
10	0,7	3,1225E-01	3,1006E-01	3,0785E-01	3,0563E-01	3,0339E-01	3,0114E-01	2,9887E-01	2,9659E-01	2,9431E-01	2,9200E-01
11	0,8	2,8969E-01	2,8737E-01	2,8504E-01	2,8269E-01	2,8034E-01	2,7798E-01	2,7562E-01	2,7324E-01	2,7086E-01	2,6848E-01
12	0,9	2,6609E-01	2,6369E-01	2,6129E-01	2,5888E-01	2,5647E-01	2,5406E-01	2,5164E-01	2,4923E-01	2,4681E-01	2,4439E-01
13	1,0	2,4197E-01	2,3955E-01	2,3713E-01	2,3471E-01	2,3230E-01	2,2988E-01	2,2747E-01	2,2506E-01	2,2265E-01	2,2025E-01
14	1,1	2,1785E-01	2,1546E-01	2,1307E-01	2,1069E-01	2,0831E-01	2,0594E-01	2,0357E-01	2,0121E-01	1,9886E-01	1,9652E-01
15	1,2	1,9419E-01	1,9186E-01	1,8954E-01	1,8724E-01	1,8494E-01	1,8265E-01	1,8037E-01	1,7810E-01	1,7585E-01	1,7360E-01
16	1,3	1,7137E-01	1,6915E-01	1,6694E-01	1,6474E-01	1,6256E-01	1,6038E-01	1,5822E-01	1,5608E-01	1,5395E-01	1,5183E-01
17	1,4	1,4973E-01	1,4764E-01	1,4556E-01	1,4350E-01	1,4146E-01	1,3943E-01	1,3742E-01	1,3542E-01	1,3344E-01	1,3147E-01
18	1,5	1,2952E-01	1,2758E-01	1,2566E-01	1,2376E-01	1,2188E-01	1,2001E-01	1,1816E-01	1,1632E-01	1,1450E-01	1,1270E-01
19	1,6	1,1092E-01	1,0915E-01	1,0741E-01	1,0567E-01	1,0396E-01	1,0226E-01	1,0059E-01	9,8925E-02	9,7282E-02	9,5657E-02
20	1,7	9,4049E-02	9,2459E-02	9,0887E-02	8,9333E-02	8,7796E-02	8,6277E-02	8,4776E-02	8,3293E-02	8,1828E-02	8,0380E-02
21	1,8	7,8950E-02	7,7538E-02	7,6143E-02	7,4766E-02	7,3407E-02	7,2065E-02	7,0740E-02	6,9433E-02	6,8144E-02	6,6871E-02
22	1,9	6,5616E-02	6,4378E-02	6,3157E-02	6,1952E-02	6,0765E-02	5,9595E-02	5,8441E-02	5,7304E-02	5,6183E-02	5,5079E-02
23	2,0	5,3991E-02	5,2919E-02	5,1864E-02	5,0824E-02	4,9800E-02	4,8792E-02	4,7800E-02	4,6823E-02	4,5861E-02	4,4915E-02
24	2,1	4,3984E-02	4,3067E-02	4,2166E-02	4,1280E-02	4,0408E-02	3,9550E-02	3,8707E-02	3,7878E-02	3,7063E-02	3,6262E-02
25	2,2	3,5475E-02	3,4701E-02	3,3941E-02	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2,9655E-02	2,8985E-02

Essendo la Gaussiana simmetrica indichiamo solamente la parte positiva perché la parte negativa viene semplicemente specchiata.
Questi valori si trovano come l'incrocio di una riga e di una colonna, la funzione che viene usata su Excel si chiama DISTRIB.NORM.ST.N([valore della funzione], FALSO)

- questo "FALSO" che vuol dire "calcola la densità di probabilità
- se fosse scritto "VERO" vorrebbe dire "calcola la probabilità cumulata"

Se vogliamo sapere un certo percentile (quello che mi lascia il 95% a sinistra che per simmetria si ottiene trovando il valore che lascia il 2,5% a destra).

Esercizio A1 (su Excel)

Il valore della rigidezza a compressione di una molla ad elica cilindrica in acciaio segue una distribuzione normale con media $2000\frac{N}{m}$ e scarto quadratico medio di $20\frac{N}{m}$.

a) Calcolare la percentuale di molle la cui rigidezza sarà superiore a $1950\frac{N}{m}$

   Si può calcolare in due modi:

   1. Calcolo innanzitutto la posizione percentuale sulla gaussiana corrispondente a y come $\frac{1950-m}{s} = -2{,}5\%$
      Poi calcoliamo la probabilità di essere maggiori di 1950 la calcoliamo come
      $1 - DISTRIB.NORM - ST(-2{,}5\%) = 99{,}38$
   2. Facciamo tutto in un unico passaggio con
      $1 - DISTRIB.NORM(1950, m, s; VERO) = 99{,}38\%$

b) Calcolare la percentuale di molle la cui rigidezza sarà inferiore a $2040\frac{N}{m}$
   Si fa lo stesso procedimento del caso a) e si ottiene 2% e 97,72%

c) Se la tolleranza ammessa a disegno per la rigidezza delle molle è di $30\frac{N}{m}$ indicare la percentuale di molle in tolleranza
   Facciamo il valore medio + e – 30 e otteniamo le variabili standardizzate e le corrispettive percentuali 1,5% e -1,5%. Poi facciamo la differenza fra le due normali
   $DISTRIB.NORM.ST(1.5) - DISTRIB.NORM.ST(-1.5) = 86.64\%$