Statistica I, Laurea triennale in Ing. Gestionale, a.a. 20010/11
Registro delle lezioni

Lezione 1 (27/10). Introduzione al corso; materiale e comunicazioni alla pag. di F. Flandoli:

http://www2.ing.unipi.it/~a008484/dispStatisticaGestionaleTriennale.html.

Si osserva che ci sono tre corsi a carattere statistico nel percorso di Gestionale a Pisa, ovvero Statistica I alla triennale, Statistica II alla magistrale, Statistica Applicata (Prof. Lanzetta) tra i facoltativi della triennale, corsi coordinati ta loro.

La statistica si può grosso modo dividere in Statistica descrittiva da un lato, e Statistica basata sul Calcolo delle Probabilità dall'altro. Il capitolo 2 del Ross è dedicato al primo argomento. Verrà trattato separatamente. Gli altri capitoli puntano alla Statistica basata sul Calcolo delle Probabilità.

Si inizia lo studio del capitolo 3, sui fondamenti del calcolo delle probabilità. Vengono date le definizioni di universo, eventi, eventi elementari o esiti, vengono discusse le operazioni (unione, intersezione e complementare) su eventi e viene poi definita la probabilità (interpretata intuitivamente col concetto di massa). Vengono illustrazione alcune regole regole, su $P(A^c)$ e $P(A \cup B)$ anche nel caso non disgiunto (interpretate tramite l'idea di massa; la prima anche dimostrata).

Sono poi descritti gli spazi a esiti equiprobabili, la regola per calcolare $P(A)$ (rapporto tra casi favorevoli e casi possibili) e viene così calcolata la probabilità che esca almeno un 6 nel lancio di due dadi.

Lezione 2 (30/10). Principio di enumerazione (anche generalizzato) ed esempio 3.5.1. Permutazioni ed esempi 3.5.2. Disposizioni e combinazioni, illustrate col problema di creare sequenze oppure gruppi di lettere (diverse) di un alfabeto: le sequenze ordinate di k lettere diverse prese da un alfabeto di n lettere, sono le disposizioni di n oggetti in k posti, $n \cdot (n-1) \cdot ... \cdot (n-k+1)$; invece i gruppi (non ordinati) di k lettere diverse prese da un alfabeto di n lettere, sono le combinazioni di n oggetti in k posti, $\binom{n}{k} = \frac{n \cdot (n-1) \cdot ... \cdot (n-k+1)}{k!}$ (le prime differiscono per l'ordinamento). Combinazioni e stringhe di 0/1: il numero di stringhe lunghe n con k uni è $\binom{n}{k}$; riconoscimento dell'equivalenza tra i due problemi (gruppi di lettere e stringhe). Esempio 3.5.4 (detta distribuzione *ipergeometrica*).

Probabilità condizionata, interpretazione grafica. Formula di fattorizzazione (probabilità totali) ed esempio 3.7.1. Formula di Bayes ed esempio 3.7.2.

Viene interpretata la formula delle probabilità totali tramite un albero di eventi. Si tratta di calcolare probabilità lungo i percorsi e sommarle.

Esercizi per casa su fattorizzazione e Bayes: es. 10 (H2) del 28/05/2010; es. 1 del 20/07/2010; es. 1.i del 29/06/2010. Si consiglia anche la lettura degli esempi 3.5.3, 3.7.4, 3.7.7.

Lezione 3 (3/11). Vedere la registrazione di Barsanti. Introduzione alla statistica, statistica descrittiva. Frequenza assoluta e relativa. Modi di rappresentare graficamente i dati. Aggregazione di dati in classi.

Campione, media campionaria e sue proprietà. Media pesata. Mediana e moda.

Lezione 4 (6/11). La formula di Bayes risolve il problema di trovare la causa più probabile, osservato l'effetto (mentre il concetto di probabilità condizionale parla della probabilità dell'effetto data la causa). Se l'effetto A può essere causato da B_1 o B_2, si vuole capire chi è maggiore tra $P(B_1|A)$ e $P(B_2|A)$. Siccome il denominatore (nella formula di Bayes) è lo stesso, basta confrontare $P(A|B_1)P(B_1)$ con $P(A|B_2)P(B_2)$. Nel caso di cause equiprobabili, si confronta $P(A|B_1)$ con $P(A|B_2)$ (intuitivo). E' quanto avviene nei canali con rumore, quando il ricevente deve decidere il messaggio inviato (se B_1 o B_2), sulla base della ricezione di un messaggio corrotto A.

Viene interpretato, con l'albero di eventi, il problema della scelta della causa più probabile. Si tratta di calcolare probabilità lungo percorsi e confrontarle.

Formule di fattorizzazione e di Bayes in generale. **Esercizio**: in un sistema di lettura automatica dei testi scritti a mano, le lettere n ed u sono facili da confondere. In una certa lingua, la n compare con frequenza 1/15, mentre la u con frequenza 1/25. Se la lettera è davvero n, il sistema legge n il 90% delle volte. Se la lettera è u, il sistema la legge n il 20% delle volte. Se la lettera è diversa sia da n sia da u, il sistema la legge n il 2% delle volte. Se il sistema legge n, con che probabilità ha sbagliato? Soluzione: 0.302.

Per casa: a titolo di esercizio, risolvere i problemi esposti negli esempi 3.7.4 e 3.7.7.

Indipendenza tra eventi. Esempio 3.8.4.

Variabili aleatorie (esempi). V.a. discrete e continue, densità discreta (funzione massa di probabilità) e densità continua.

Lezione 5 (10/11). Vedere la registrazione di Barsanti. Varianza e deviazione standard campionaria con esempi e tecniche per abbreviare il calcolo. Percentili campionari e quantili. Box plot.

Formula di Chebyshev con dimostrazione ed esempi di applicazioni. Campioni normali, asimmetrici, bimodali.

Lezione 6 (13/11). Capitolo 4, paragrafi 1 e 2. Variabili aleatorie discrete e continue. Densità discreta, $p(k) = P(X = k)$, sue proprietà, $P(X \in A) = \sum_{k \in A} p(k)$. Densità continua $f(x)$, sue proprietà, $P(X \in A) = \int_A f(x)\,dx$. I numeri $P(X = a)$ determinano tutto nel caso discreto, valgono 0 nel caso continuo. Grafici nei due casi.

Funzione di distribuzione cumulativa ($F(t) = P(X \leq t)$).

Capitolo 5, paragrafo 1. V.a. di Bernoulli e binomiale. Grafici. Ammissibilità delle formula binomiale (ovvero verifica che sono numeri non negativi a somma 1).

Teorema: la somma $S = X_1 + ... + X_n$ di n Bernoulli indipendenti di parametro p è una $B(n, p)$. Dimostrazione. Interpretazione di S come numero di successi in n prove.

Esempio: numero di correntisti di una banca che prelevano (ipotesi semplificate). Calcolo di un numero k tale che $P(S > k) = 0.99$ (impostato teoricamente).

Esercizi per casa: risolvere i problemi enunciati agli esempi 5.1.1, 5.1.3 punto (a).

Lezione 7 (17/11). Soluzione di alcuni esercizi assegnati per casa nelle lezioni scorse.

Lezione 8 (20/11). Valori medi e loro proprietà. Valor medio (atteso) di una v.a. discreta, valor medio di una v.a. continua. Somiglianza con la media empirica di un campione: detti a_k i valori della v.a. X e $p(a_k)$ le loro probabilità, la def. di media è $E[X] = \sum a_k p(a_k)$, mentre vale

$$\overline{x} = \frac{x_1 + ... + x_n}{n} = \sum a_k \widehat{p}_k$$

dove $\widehat{p}_k$ è la frequenza relativa empirica con cui si osserva a_k nel campione. Interpretazione grafica. Calcolo di $E[X]$ in alcuni esempi, incluse Bernoulli e uniforme su $[0, 1]$.

Valor medio di una trasformazione di v.a., $E[g(X)]$, esempi. Caso particolare: momenti e varianza. Deviazione standard σ, interpretazione grafica; misura di incertezza. Calcolo della varianza per le Bernoulli.

Proprietà del valor medio e della varianza:

$$E[aX + bY + c] = aE[X] + bE[Y] + c$$
$$Var[aX + c] = a^2 Var[X]$$

e, se X ed Y sono *indipendenti*,

$$Var[X+Y] = Var[X] + Var[Y].$$

Applicazione al calcolo di media e varianza delle binomiali, attraverso il teorema della lezione 6.

Esempio grafico di μ e σ per una $B(n,p)$ con n elevato e p piccolo, con applicazione al dimensionamento di un magazzino, relativamente a prodotti acquistati di rado, pur in presenza di moltissimi potenziali clienti.

Lezione del 24/11: non tenuta per occupazione aule polo F.

Lezione 9 (27/11). Funzione generatrice dei momenti, def. generale e formule nel caso discreto e continuo. Legame coi momenti. Teorema sulla generatrice della somma di v.a. indipendenti. Esempi: Bernoulli e binomiale.

V.a. di Poisson: definizione, verifica della somma pari ad 1, calcolo della generatrice e da essa di media e varianza. Teorema degli eventi rari (convergenza della binomiale alla Poisson) con dimostrazione. Verifica non rigorosa delle formule per la generatrice, la media e la varianza, usando questa convergenza. Grafico tipico di una Poisson. Importanza applicativa della Poisson: dipende da un solo parametro, pari al valor medio e quindi approssimabile sperimentalmente in modo facile, a differenza delle binomiali; facile quindi da usare in applicazioni come i sistemi di servizio (magazzini ecc.) con molti utenti potenziali, ciascuno con bassa probabilità dichiedere il servizio.

V.a. esponenziali: definizione, verifica dell'area pari ad 1, calcolo della generatrice e da essa di media e varianza. Nota: $\sigma = \mu$, elevata aleatorietà rispetto a binomiali e Poisson (dove $\sigma = \sqrt{\mu}$). Interpretazione grafica di μ e σ per le esponenziali (casi λ grande e λ piccolo).

Lezione 10 (1/12). Vedere la registrazione di Barsanti. Variabili aleatorie Gaussiane e loro funzione generatrice. Media e varianza di una variabile Gaussiana. Trasformazione lineare di una variabile Gaussiana e standardizzazione. Esercizi su v.a. Gaussiane. Riproducibilita' delle v.a. Gaussiane con esempio.

Lezione 11 (4/12). Capitolo 6. Concetto di campione $X_1, ..., X_n$. Media $\overline{X}$ e varianza campionaria S^2 (come variabili aleatorie). Stimatori corretti, verifica che la media campionaria è uno stimatore corretto di μ. Enunciato del fatto che la varianza campionaria è uno stimatore corretto di σ^2 e del fatto che $\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$ è uno stimatore corretto di σ^2; deduzione della formula $E\left[\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2\right] = \frac{n-1}{n}\sigma^2$ (cioè $\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$ non stima correttamente σ^2).

Calcolo di $Var\left[\overline{X}\right]$ $(= \frac{\sigma^2}{n})$, grafico indicativo della densità di $\overline{X}$, indicazione della sua vicinanza a μ.

Esercizi dai compiti d'esame. Osservazioni sulla proprietà di autoriproduzione: i) le gaussiane sono autoriproducenti, nel senso più ampio possibile (lezione precedente: se X ed Y sono gaussiane indipendenti, allora $aX+bY+c$ è gaussiana); ii) le binomilai lo sono in un senso un po' ristretto: la somma di una $B(n,p)$ più una $B(k,p)$ indipendenti è $B(n+k,p)$ (si verifica col teorema di legame tra binomiali e Bernoulli); iii) per le Poisson, vale che la somma di una $\mathcal{P}(\lambda)$ più una $\mathcal{P}(\mu)$ indipendenti è $\mathcal{P}(\lambda+\mu)$ (verificato con le funzioni generatrici, come per le gaussiane).

Lezione 12 (11/12). Standardizzazione di una v.a. e della somma $S_n = X_1 + ... + X_n$ di un campione. Proprietà nel caso gaussiano. Teorema limite centrale e sua applicazione negli esercizi. Esercizi dai compiti d'esame.

Distribuzione della media campionaria: esatta nel caso di variabili gaussiane, approssimata nel caso generale. Dimostrazione del fatto che S^2 è uno stimatore corretto ($E[S^2] = \sigma^2$). Non è stato svolto il collegamento con chi quadro e t di Student.

Esercizi suggeriti per il compitino:

tutti quelli svolti a lezione/esercitazione;

28/05/2010: domande 2,3,4,6,7 (alcune domande, come la 7 di questo esercizio, richiedono la lezione dell'11);

8/06/2010: es. 1.iii, es. 2 tutto;

29/06/2010: es. 1 tutto, es. 2.ii;

20/07/2010: es. 1, i, ii, iii, iv, vi, vii;

14/09/2010: es 1 tutto, es. 2.ii.

Facoltativamente si possono anche prendere esercizi dai compiti degli anni ancora precedenti, selezionando quelli simili (operazione molto utile per lo studente).

Lezione 13 (15/12). Vedere la registrazione di Barsanti. Stima puntuale usando il principio di massima verosimiglianza. Esempi: stima del parametro della legge esponenziale e stima dei parametri della legge gaussiana.

Stima di intervalli. Definizione di confidenza e intervallo di confidenza. Stima di intervalli di confidenza bilaterali e unilaterali per la media di una v.a. gaussiana di cui è supposta nota la sigma.

Lezione 14 (18/12). A causa delle numerose assenze, la lezione viene riportata qui integralmente. Presenta un problema pratico di stima (che verrà proseguito nelle prossime lezioni).

Un'azienda che effettua interventi e riparazioni vuole stimare due grandezze, per poter dimensionare l'organico ed organizzare i turni di lavoro. La prima grandezza è il numero medio μ di ore di lavoro in azienda, giornaliere, necessarie per effettuare tutti i lavori richiesti. La seconda è la probabilità p di dover effettuare interventi esterni. Indichiamo con N il numero di ore di lavoro interne, con X una v.a. di Bernoulli che vale 1 se c'è da effettuare un lavoro esterno (entrambe le variabili sono riferite ad una giornata, generica).

L'azienda si pone le seguenti domande: i) come stimare μ e p? ii) Che errore potremmo aver commesso in tale stima? iii) Quante osservazioni servono per fare tali stime?

Supponendo di avere a che fare con un'azienda di media grandezza, in cui i valori di N siano di varie decine e non di pochissime unità, decidiamo di trattare N come una v.a. continua e per semplicità gaussiana. Invece X è intrinsecamente Bernoulli. Dobbiamo stimare in entrambi i casi il valor medio:

$$\mu = E[N], \qquad p = E[X].$$

La risposta alla domanda (i) in un certo senso è ovvia: si devono effettuare n rilevazioni giornaliere delle due grandezze, chiamiamole

$$N_1, ..., N_n \qquad \text{e} \qquad X_1, ..., X_n$$

(anche numerosità diverse per i due problemi) e poi calcolare gli stimatori

$$\widehat{\mu} = \frac{N_1 + ... + N_n}{n}, \qquad \widehat{p} = \frac{X_1 + ... + X_n}{n}.$$

Detto questo però sorgono tante domande, appunto ad esempio le domande (ii) ed (iii), circa la bontà di queste stime.

Avendo ipotizzato che N è gaussiana, vale

$$\mu = \widehat{\mu} \pm \frac{\sigma q_{1-\frac{\alpha}{2}}}{\sqrt{n}}$$

con confidenza $1 - \alpha$. Ad esempio, $\mu = \widehat{\mu} \pm \frac{\sigma \cdot 1.96}{\sqrt{n}}$ al 95%. Questo significa che, al 95%, il massimo errore possibile è $\frac{\sigma \cdot 1.96}{\sqrt{n}}$. (In particolare, non c'è un errore massimo possibile certo, ma sempre a meno di una piccola probabilità; c'è sempre una piccola probabilità che l'errore sia ancora più grosso). Questo non significa che l'errore sarà pari a $\frac{\sigma \cdot 1.96}{\sqrt{n}}$, al 95%: *al massimo* sarà $\frac{\sigma \cdot 1.96}{\sqrt{n}}$.

Ma se riduciamo la confidenza, esso è minore:

$$\begin{array}{rl} \text{al } 90\%: & \frac{\sigma \cdot 1.64}{\sqrt{n}} \\ \text{all' } 80\%: & \frac{\sigma \cdot 1.28}{\sqrt{n}} \\ \text{al } 70\%: & \frac{\sigma \cdot 1.04}{\sqrt{n}} \\ \text{al } 60\%: & \frac{\sigma \cdot 0.84}{\sqrt{n}} \end{array}$$

e così via. L'idea si vede bene graficamente tracciando la densità gaussiana di $\frac{N_1+...+N_n}{n}$ ed osservando come varia l'intervallo attorno a μ quando si varia l'area soprastante. Quindi è molto probabile che l'errore sia molto più piccolo di $\frac{\sigma \cdot 1.96}{\sqrt{n}}$, ad esempio sia la metà. Il numero $\frac{\sigma \cdot 1.96}{\sqrt{n}}$ fornisce l'ordine di grandezza.

Veniamo ora all'aspetto pratico: supponiamo di aver fatto $n = 25$ osservazioni ed aver trovato $\widehat{\mu} = 62.8$. Che possiamo dire, ad esempio al 95%? Che

$$\mu = 62.8 \pm \frac{\sigma \cdot 1.96}{5} = 62.8 \pm 0.39 \cdot \sigma.$$

Ma quanto vale σ? Nessuno ce lo può dire. La cosa più naturale, avendo a disposizione il campione di numerosità $n = 25$, è calcolare S. Supponiamo di farlo ed aver trovato $S = 18.3$. Allora, approssimativamente (S non è σ), possiamo affermare che al 95%

$$\mu = 62.8 \pm 0.39 \cdot 18.3 = 62.8 \pm 7.14.$$

In altre parole, al 95%, il valore incognito μ è compreso tra 55.66 e 69.94. Ma, come detto sopra, molto probabilmente è abbastanza più vicino a 62.8. Ad esempio, al 60%, vale circa

$$\mu = 62.8 \pm 3.5$$

cioè μ è compreso tra 59.3 e 66.3.

la sostituzione di σ con S ha introdotto un'approssimazione. Un teorema dice che i risultato (cioè l'ampiezza dell'intervallo di confidenza) torna ad essere un risultato esatto se, oltre a sostituire σ con S, si sostituisce il quantile gaussiano standard $q_{1-\frac{\alpha}{2}}$ con il quantile della t di Student a $n-1$ gradi di libertà:

$$\mu = \widehat{\mu} \pm \frac{\sigma t^{(n-1)}_{1-\frac{\alpha}{2}}}{\sqrt{n}}.$$

Nel nostro esempio, usando le tavole, vale $t^{(24)}_{1-\frac{0.05}{2}} = 2.064$ e quindi

$$\mu = 62.8 \pm \frac{18.3 \cdot 2.064}{5} = 62.8 \pm 7.55.$$

Il risultato è un po' peggiore di quello approssimato precedente, ma è sicuro. La differenza non è però marcatissima.

La domanda (ii) ha però una variante fondamentale: che si parli di errore relativo invece che assoluto. L'errore assoluto è $|\widehat{\mu} - \mu|$ mentre l'errore relativo è

$$\left| \frac{\widehat{\mu} - \mu}{\mu} \right|.$$

Allora l'errore relativo massimo possibile con confidenza $1 - \alpha$ è

$$\left| \frac{\widehat{\mu} - \mu}{\mu} \right| = \frac{\sigma q_{1-\frac{\alpha}{2}}}{\sqrt{n}\,|\mu|}.$$

Nel nostro esempio, al 95%, usando ad esempio per semplicità i quantili guassiani

$$\left| \frac{\widehat{\mu} - \mu}{\mu} \right| = \frac{18.3 \cdot 1.96}{5 \cdot |\mu|} = \frac{7.17}{|\mu|}.$$

naturalmente nessuno ci può dare μ, visto che è la quantità da stimare. Quindi approssimativamente sostiuiamola con $\widehat{\mu}$ che è nota:

$$\left| \frac{\widehat{\mu} - \mu}{\mu} \right| \approx \frac{7.17}{62.8} = 0.114.$$

In sostanza, si commette un errore relativo di un decimo (decente per scopi di commercio non troppo spinti). Ovviamente se si vuole usare la t di Student, viene lievemente più grande (provare).

Sempre relativamente a N, veniamo alla domanda (iii). Il numero di osservazioni da fare non può essere una grandezza assoluta, indipendente da requisiti. Dipende dalla precisione che vogliamo ottenere e dalla confidenza che scegliamo (il rischio che accettiamo di correre).

Lezione 15 (22/12). (Proseguimento dell'esercizio sulla stima iniziato la volta precedente.)

Sempre relativamente alla v.a. N della lezione precedente, veniamo alla domanda (iii). Essa è un esempio di DOE (*Design Of Experiments*).

Il numero di osservazioni da fare non può essere una grandezza assoluta, indipendente da requisiti. Dipende dalla precisione che vogliamo ottenere e dalla confidenza che scegliamo (il rischio che accettiamo di correre; rischio di fare una dichiarazione falsa circa l'intervallo in cui cade la media). Supponiamo di correre un rischio del 5%, prendere cioè confidenza 95% e supponiamo di volere un errore (massimo) pari a 5, erroe assoluto. Uguagliando l'errore massimo a 5 abbiamo $\frac{\sigma \cdot 1.96}{\sqrt{n}} = 5$, ovvero

$$n = \left(\frac{\sigma \cdot 1.96}{5} \right)^2 = 0.154 \cdot \sigma^2.$$

Con l'uguaglianza si intende in realtà il primo intero $n \geq 0.154 \cdot \sigma^2$ (infatti per essere più precisi andrebbe impostata dall'inizio la disuguaglianza $\frac{\sigma \cdot 1.96}{\sqrt{n}} \leq 5$). Resta il grave problema di conoscere σ: se non abbiamo ancora fatto rilevazioni, se non abbiamo dati, σ è incognita. Non ci sono scappatoie generali: o si conosce un valore approssimato di σ sulla base di dati precedenti, oppure si deve ipotizzare l'ordine di grandezza di σ, approssimando ovviamente per eccesso. Senza σ non si può stabilire n in anticipo. Se non si hanno dati precedenti o capacità di stima dell'ordine di grandezza, bisogna iniziare i campionamenti, raccogliere un po' di dati e con essi stimare σ. Questi primi dati concorreranno comunque alla stima finale di μ. Supponiamo di aver raccolto una decina di dati preliminari, dai quali esca la stima

$$S = 20.4$$

Allora troviamo

$$n = 0.154 \cdot 20.4^2 = 64.089.$$

Servono circa 65 osservazioni. In realtà, dopo un po' di ulteriori osservazioni conviene ri-stimare σ per rendere più accurata la previsione del numero di osservazioni da fare.

Se volevamo invece l'errore relativo (massimo) assegnato, es. 10%, dovevamo imporre

$$\frac{\sigma \cdot 1.96}{\sqrt{n}\,|\mu|} = 0.1$$

ovvero

$$n = \left(\frac{\sigma \cdot 1.96}{0.1 \cdot |\mu|} \right)^2 = 384.16 \cdot \left(\frac{\sigma}{|\mu|} \right)^2.$$

Qui servono addirittura una stima preliminare di σ e μ. Si agisce come sopra. Supponiamo che dopo alcune osservazioni preliminari abbiamo trovato

$\overline{x} = 60.5$, $S = 20.4$. Allora

$$n = 384.16 \cdot \left(\frac{20.4}{60.5}\right)^2 = 43.678.$$

Questi esempi numerici hanno il solo scopo di vedere le cose fino in fondo e vedere la ragionevolezza del risultato.

Si noti comunque che questi calcoli producono valori piuttosto alti di n. In certe applicazioni pratiche, molte decine di osservazioni sono davvero costose. C'è un rimedio? Ricordiamo quanto appreso sopra circa l'intervallo di confidenza: esso esprime il risultato più pessimistico. Con buona probabilità, l'intervallo al 95% è pessimistico, la stima è molto migliore, come evidenzia l'intervallo al 60%, ad esempio.

Se accettassimo un rischio molto alto, 40%, i calcoli precedenti darebbero:

$$n^{60\%}_{assoluto} = \left(\frac{\sigma \cdot 0.84}{5}\right)^2 = 0.028 \cdot \sigma^2 \overset{S=20.4}{=} 0.028 \cdot 20.4^2 = 11.652.$$

Naturalmente non possiamo esporci ad un tale rischio, ma questo calcolo ci dice che il 60% delle volte accadrebbe che 12 osservazioni sono sufficienti, invece che 65. Similmente, accettando un rischio del 20%,

$$n^{80\%}_{assoluto} = \left(\frac{\sigma \cdot 1.28}{5}\right)^2 = 0.065 \cdot \sigma^2 \overset{S=20.4}{=} 0.065 \cdot 20.4^2 = 27.05.$$

Insomma, con elevata probabilità, bastano molte meno osservazioni. Che fare? Ovviamente si può decidere di fare poche osservazioni (es. solo 20-30) e sperare che le cose siano andate bene. Si può però tracciare un grafico della stima della media $\widehat{\mu}$ al crescere del numero di prove. Nel senso: dopo aver eseguito n osservazioni, calcoliamo $\widehat{\mu}_n$ ed aggiungiamolo al grafico precedentemente fatto dei valori di $\widehat{\mu}$ in funzione del numero di prove. Al crescere di n questo grafico tenderà ad assestarsi attorno all'asintoto orizzontale μ (asintoto però sconosciuto!). Quando vediamo il grafico diventare sufficientemente orizzontale, abbiamo un forte sintomo che siamo già "arrivati a convergenza", come si suol dire. Non c'è la certezza assoluta, ma è molto difficile che un tale grafico si assesti e poi riprenda a muoversi. Bene, nel 60% dei casi, si assesta molto presto, nell'80%, poco oltre, e così via. Solo in rari casi necessita davvero di valori di n intorno a 65 per assestarsi; è solo il caso più pessimistico, che però è garantito al 95%. A priori, non possiamo sapere se ci capiterà questo caso o quelli più fortunati. Bisogna eseguire le prove sequenzialmente e sperare. Quanto qui espresso è una versione pratica della cosidetta *Sequential Analysis*.

Ripetiamo ora alcuni dei passi precedenti per il problema della stima della proporzione p, altro problema classico e ricorrente. Lo stimatore è $\widehat{p}$, ma ora non vale più la teoria gaussiana dell'intervallo di confidenza. Tuttavia, in modo approssimato essa è ancora vera: vale

$$p = \widehat{p} \pm \frac{\sigma q_{1-\frac{\alpha}{2}}}{\sqrt{n}}, \qquad \sigma^2 = Var\left[X\right] = p\left(1-p\right)$$

con confidenza approssimativamente pari $1-\alpha$. Ciò che è approssimata è la probabilità che p stia nell'intervallo suddetto, non l'intervallo in sé. Tutto deriva dal teorema limite centrale, in quanto

$$\begin{aligned} P\left(\left|\widehat{p}-p\right| \leq \frac{\sigma q_{1-\frac{\alpha}{2}}}{\sqrt{n}}\right) &= P\left(\left|\frac{X_1+...+X_n}{n}-p\right| \leq \frac{\sigma q_{1-\frac{\alpha}{2}}}{\sqrt{n}}\right) \\ &= P\left(\left|\frac{X_1+...+X_n-np}{\sqrt{n}\sigma}\right| \leq q_{1-\frac{\alpha}{2}}\right) \approx 1-\alpha \end{aligned}$$

dove l'ultima approssimazione è fornita appunto dal TLC.

Facciamo un esempio pratico: supponiamo di aver fatto $n = 25$ osservazioni ed aver trovato $\widehat{p} = 0.21$. Che possiamo dire, ad esempio al 95%? Che con probabilità circa uguale a questa, vale

$$p = 0.21 \pm 0.39 \cdot \sigma$$

un po' come nel caso gaussiano. Resta il problema di conoscere σ.

Qui però c'è un elemento in più, molto particolare: $\sigma^2 = p\left(1-p\right)$. Il parametro σ è legato alla quantità p che stiamo cercando di stimare. Una prima conclusione quindi è che valga, approssimativamente

$$p = \widehat{p} \pm \frac{\sqrt{\widehat{p}\left(1-\widehat{p}\right)}q_{1-\frac{\alpha}{2}}}{\sqrt{n}}.$$

Nel nostro esempio,

$$p = 0.21 \pm 0.39 \cdot \sqrt{0.21 \cdot (1-0.21)} = 0.21 \pm 0.16.$$

Vale cioè

$$0.05 \leq p \leq 0.37.$$

Non è un risultato eccellente, in senso relativo. Naturalmente, è abbastanza probabile che l'intervallo sia più piccolo, come abbiamo visto nel caso gaussiano: ad esempio, all'80% vale

$$p = 0.21 \pm \frac{1.28}{5} \cdot \sqrt{0.21 \cdot (1 - 0.21)} = 0.21 \pm 0.104.$$

cioè p è compreso tra 0.1 e 0.3. Parlando a braccio, la frequenza con cui il negozio deve mandare operatori fuori sede si aggira tra 1/10 e 3/10. Se questa vaghezza di informazione è sufficiente, basta così, altrimenti bisogna campionare di più.

L'errore relativo in astratto è

$$\left| \frac{\widehat{p} - p}{p} \right| \leq \frac{\sqrt{p\,(1-p)}\,q_{1-\frac{\alpha}{2}}}{p\sqrt{n}} = \frac{\sqrt{\frac{1-p}{p}}\,q_{1-\frac{\alpha}{2}}}{\sqrt{n}}$$

ed approssimando le espressioni sulla destra diventa

$$\left| \frac{\widehat{p} - p}{p} \right| \leq \frac{\sqrt{\frac{1-\widehat{p}}{\widehat{p}}}\,q_{1-\frac{\alpha}{2}}}{\sqrt{n}}$$

in questo esempio è (approssimativamente al 95%)

$$\left| \frac{\widehat{p} - p}{p} \right| \leq \frac{\sqrt{\frac{1-0.21}{0.21}}\,1.96}{5} = 0.76.$$

Per certe applicazioni è davvero troppo grosso, per altre può anche essere accettabile.

Si deve notare che è venuto così grosso perche' $\widehat{p}$ è piccolo: se si stima una proporzione piccola, la tendenza è di commettere un errore relativo grosso. Se invece $\widehat{p}$ fosse stato grande, $\sqrt{\frac{1-\widehat{p}}{\widehat{p}}}$ era piccolo e avrebbe contribuito a diminuire l'errore relativo.

Spesso nelle applicazioni si cerca di stimare una proporzione piccola al solo scopo di sapere che è piccola, non di conoscerne con precisione il valore. Sapere che è 0.05 o 0.1 o 0.15 non cambia le nostre azioni successive, anche se questi numeri differiscono di tantissimo in senso relativo. Differiscono poco in senso assoluto. Allora, in problemi di questo genere, basta chiedere che l'errore assoluto sia piccolo. L'errore relativo non serve. In sintesi, in problemi in cui basta scoprire che p è piccolo basta desiderare che l'errore assoluto sia piccolo; e quindi i difetti suddetti dell'errore relativo per $\widehat{p}$ piccolo diventano inessenziali.

In quest'ottica, immaginiamo di voler stimare p con precisione assoluta 0.1 (se $\widehat{p}$ è piccolo, ci basta, p non supererà $\widehat{p}+0.1$; se $\widehat{p}$ è grande, un errore assoluto di 0.1 non è così grave). Dobbiamo imporre

$$\frac{\sqrt{p\,(1-p)}\cdot 1.96}{\sqrt{n}}=0.1$$

ovvero

$$n=\left(\frac{1.96}{0.1}\right)^2 p\,(1-p)\,.$$

Serve una stima di p, che in fase di DOE può provenire da campionamenti precedenti, da primi piccoli campionamenti, da ipotesi. Ma in questo caso vale anche la seguente stima universale: siccome l'espressione $p\,(1-p)$ può al massimo valere $\frac{1}{4}$, Alla peggio dovremo prendere

$$n=\left(\frac{1.96}{0.1}\right)^2\frac{1}{4}=96.04.$$

Ovviamente non è un valore molto incoraggiante, però è universale. E' chiaro che all'80% basta

$$n=\left(\frac{1.28}{0.1}\right)^2\frac{1}{4}=40.96$$

ed al 60% addirittura

$$n=\left(\frac{0.84}{0.1}\right)^2\frac{1}{4}=17.64.$$

Quindi, eseguendo le cose sequenzialmente e sperando di non essere troppo sfortunati, dovrebbe bastare un numero contenuto di osservazioni.

Lezione 16 (12/1). Vedere la registrazione di Barsanti. Somma di quadrati di variabili normali standard e legge del chi quadro. Riproducibilità di tale legge. Legge congiunta di media e varianza campionaria.

Intervallo di confidenza per la varianza. Covarianza e coefficiente di correlazione.

Lezione 17 (15/1). Viene introdotta la teoria dei test statistici, attraverso un esempio (compito di giugno 2010), in cui si conoscono media e varianza delle vendite settimanali in regime normale, si introduce un'attività pubblicitaria e se ne vuole capire l'effetto tramite un campione (si suppone che la varianza non sia cambiata, ma si dubita della media). Si noti che il valore μ_0 e quello sperimentale $\overline{x}$ differiscono, ma questo non basta a dire che c'è stato un cambiamento; a intuito, devono giocare un ruolo anche la deviazione standard e la numerosità campionaria. Il problema viene risolto facendo riferimento a idee già note nel corso, gli intervalli di confidenza: fissato un rischio α, se il campione provenisse dalla vecchia distribuzione, la differenza $\left|\overline{X} - \mu_0\right|$ sarebbe inferiore a $\delta = \frac{\sigma q_{1-\frac{\alpha}{2}}}{\sqrt{n}}$ con probabilità $1 - \alpha$; si controlla allora se il valore speriementale $\overline{x}$ soddisfa questa condizione, $|\overline{x} - \mu_0| < \delta$ (nel qual caso non abbiamo trovato contraddizioni tra il campione e la media ipotizzata μ_0), oppure quella opposta $|\overline{x} - \mu_0| > \delta$ (nel qual caso riteniamo non più valida la media μ_0 in quanto il campione è in contraddizione con essa).

Il ragionamento viene illustrato anche graficamente, tracciando la densità di $\overline{X}$, centrata in μ_0, di deviazione $\frac{\sigma}{\sqrt{n}}$, in cui, fissato il rischio α, si tagliano due code di area $_{1-\frac{\alpha}{2}}$ individuando così un intervallo. Se $\overline{x}$ cade in esso, è un valore ragionevole di $\overline{X}$, altrimenti no.

Il test bilaterale per la media di una gaussiana (varianza nota) appena visto viene poi riformulato nel modo canonico, tramite confronto di $|z|$ con $q_{1-\frac{\alpha}{2}}$.

Viene inoltre illustrato nell'esempio il fatto che cambiando α può cambiare l'esito del test. L'intervallo dei valori possibili di α è diviso in due parti da un numero, detto p-value (valore p): tutti gli α più grandi portano al rifiuto dell'ipotesi, quelli più piccoli non permettono il rifiuto.

L'algoritmo del test viene esemplificato ulteriormente col seguente esercizio. Un certo sistema di servizio (si pensi ad esempio agli sportelli di una banca) è ben dimensionato se ci sono in media 100 richieste al giorno (se sono di più bisogna aumentarlo, se sono di meno si stanno sprecando risorse). Forse il mercato è cambiato e le richieste non sono più 100 in media. Si registra un campione per 9 giorni:

$$98, 112, 103, 96, 108, 115, 102, 99, 109.$$

Al 95%, il servizio è ben dimensionato? Si supponga, sulla base di esperienze passate, che sia $\sigma = 4$. Sol:

$$z = \frac{\overline{x} - \mu_0}{\sigma}\sqrt{n} = \frac{104.2 - 100}{4}\sqrt{9} = 3.15$$

maggiore (in valore assoluto) di $q_{1-\frac{\alpha}{2}} = z_{\frac{\alpha}{2}} = 1.96$ (vale $\alpha = 0.05$, $\frac{\alpha}{2} = 0.025$, $1 - \frac{\alpha}{2} = 0.975$, $q_{0.975} = 1.96$). Il sistema non è ben dimensionato.

Nella risoluzione si è interpretato graficamente il test, tramite una gaussiana e le sue code. Si è insistito sul rischio, del 5%, necessariamente presente. E' l'area totale delle code. Se non si specifica tale rischio a priori, non ha senso confrontare un campione con una media ipotizzata μ_0. La media empirica $\overline{x}$ è sempre diversa da μ_0, per la casualità del campione e non perche' μ_0 sia falsa. Quindi il punto è capire se $\overline{x}$ dista da μ_0 in modo eccessivo oppure no. Il grado si anomalia della distanza di $\overline{x}$ da μ_0 è dato dalle code, individuate da α.

Vengono date alcune definizioni generali: ipotesi nulla, ipotesi alternativa, regione di rifiuto, regione di accettazione, significatività, valore p.

Nel test appena visto, l'ipotesi nulla $\mathcal{H}_0$ è "la media è μ_0", la regione di rifiuto è l'insieme di tutti i campioni sperimentali $(x_1, ..., x_n)$ tali che $|z| > q_{1-\frac{\alpha}{2}}$; la significatività α è il rischio di rifiutare un'ipotesi vera:

$$\alpha = P_{\mathcal{H}_0}\left(|Z| > q_{1-\frac{\alpha}{2}}\right)$$

ed il p-value, oltre ad essere *il più piccolo α per cui il test risulta significativo*, è anche la probabilità di avere valori più estremi di quello sperimentale:

$$p = P_{\mathcal{H}_0}\left(|Z| > |z|\right)$$

dove z è quello dei dati sperimentali, Z una $N(0,1)$. Si raffigura la densità $N\left(\mu_0, \frac{\sigma^2}{n}\right)$, si disegna il valore di $\overline{x}$, da esso si trovano le due code corrispondenti: il p-value è l'area di tali code. Si standardizza tale disegno. Dall'equazione

$$q_{1-\frac{p}{2}} = |z|$$

si trova la formula

$$p = 2 - 2\Phi\left(|z|\right).$$

Viene commentata la ragionevolezza di ciò che fa il test in termini grafici. Se vale l'ipotesi $\mathcal{H}_0$ (che nell'esercizio precedente è "il servizio è ben dimensionato", o più concisamente "$\mu_0 = 100$"), la v.a. $\overline{X}$ è una $N\left(\mu_0, \frac{\sigma^2}{n}\right)$. I suoi valori cadono molto raramente nelle due piccole code di area totale 0.05. Quindi, se il valore empirico $\overline{x}$ cade proprio in tali code, siamo indotti a concludere che era sbagliata l'ipotesi (perche' sotto di essa una tale cosa non doveva accadere). Quindi il test si può eseguire controllando se $\overline{x}$ cade nelle code della gaussiana $N\left(\mu_0, \frac{\sigma^2}{n}\right)$. Standardizzando, lo si può eseguire controllando se $z = \frac{\overline{x}-\mu_0}{\sigma}\sqrt{n}$ cade nelle code della gaussiana canonica, cioè controllando se

$$|z| > q_{1-\frac{\alpha}{2}}.$$

Si sottolinea che un test o rifiuta $\mathcal{H}_0$, oppure non è in grado di farlo. Non ha senso dire che il test conferma $\mathcal{H}_0$. Basta osservare che un certo valore speriementale $\overline{x}$ che non conduce a rifiuto, è compatibile non solo con μ_0 ma con altri valori medi. Non può quindi essere una prova che μ_0 è vero.

Lezione 18 (19/1). Vedere la registrazione di Barsanti. Significato di covarianza e correlazione, con esempio. Regressione lineare in una variabile. Stimatori col metodo dei minimi quadrati. Distribuzione degli stimatori nel caso di errore gaussiano, loro media e varianza. Stima della varianza a partire dai dati.

Lezione 19 (22/1). Test unilaterali per la media. Vengono introdotti tramite il seguente problema.

Esercizio. Un servizio ferroviario viene cancellato se il numero medio di passeggeri è inferiore a 20. I passeggeri che viaggiano ad una certa ora tra Firenze e Pisa vengono contati, per 10 giorni. I valori sono

$$17, 13, 9, 23, 14, 12, 16, 11, 14, 21.$$

A livello di significatività 95%, c'è ragione di cancellare quel servizio serale?

Soluzione. Si può eseguire il solito test, ma questo non è aderente al problema. Meglio ragionare dall'inizio. Se la media è 20 (caso limite), $\overline{X}$ è una $N\left(\mu_0, \frac{\sigma^2}{n}\right)$. Se i valori di $\overline{X}$ cadono nella coda destra, non c'è nulla di male: può significare che il numero medio di passeggeri è anche maggiore di 20, meglio ancora. Il problema sussiste solo se $\overline{X}$ cade nella coda sinistra, cioè assume valori troppo piccoli. Allora è solo la coda sinistra che rappresenta la situazione anomala. Prendiamo quindi una coda sinistra di ampiezza 0.05, senza coda destra. Il test consiste nel vedere se $\overline{X}$ cade in essa, ovvero, dopo standardizzazione, se

$$\frac{\overline{x}-\mu_0}{\sigma}\sqrt{n} < -q_{1-\alpha}.$$

Questo test *unilaterale* è più "intelligente" di quello solito *bilaterale*, vista la particolare domanda dell'esercizio.

Vengono esposti i test unilaterale destro e quello unilaterale sinistro, sempre esemplificandoli graficamente.

Viene poi svolto l'esercizio 1.i dell'8/6/2010, prima eseguendo il test per $\alpha = 0.05$ (viene significativo) ed $\alpha = 0.01$ (viene non significativo), capendo graficamente che c'è un valore di demarcazione, il valore p. Poi viene calcolato questo valore, in due modi: i) osservando che esso separa l'intervallo dei valori α per cui $z > q_{1-\alpha}$ (test significativo, si tratta di un unilaterale destro) da quello dei valori per cui $z < q_{1-\alpha}$ (test non significativo), quindi risolve l'equazione $z = q_{1-p}$, da cui si trova $\Phi(z) = 1 - p$ ed infine

$$p = 1 - \Phi(z);$$

ii) ricordando che il p-value è anche la probabilità che Z assuma valori più estremi di quello sperimentale, per cui (ci si aiuti con un disegno)

$$p = P(Z > z) = 1 - \Phi(z).$$

Ci si pone il problema di confrontare i p-value dei test bilaterale e unilaterale, giustificando questo interesse come analisi di quale sia il test migliore (quello che più facilmente risulta significativo). Nel caso del test unilaterale destro (il ragionamento per l'altro è analogo), se $z > 0$, vale

$$p_B = 2 - 2\Phi(|z|) = 2 - 2\Phi(z) = 2 \cdot p_U$$

cioè il p-value migliore è quello unilaterale (è la metà dell'altro).

Vengono introdotti gli errori di prima e di seconda specie, verificando in primo luogo che α, la significatività, è la probabilità dell'errore di prima specie. Viene introdotta la probabilità dell'errore di seconda specie, osservando che non è calcolabile, non essendo prescritto un valore preciso della media della gaussiana in gioco. Fissando tale media μ, diversa da μ_0 (si può pensare che si stia fissando coem ipotesi alternativa $\mathcal{H}_1$ l'ipotesi che la media sia precisamente μ), si può calcolare la probabilità dell'errore di seconda specie quando la media è μ:

$$\beta(\mu) := P_\mu\left(|Z| < q_{1-\frac{\alpha}{2}}\right).$$

Si noti che ora Z non è una gaussiana canonica (è gaussiana, ma la media non è zero, in quanto $Z = \frac{\overline{X}-\mu_0}{\sigma}\sqrt{n}$ e $\overline{X}$ ha media μ e non μ_0). Si trova

$$\beta(\mu) = \Phi\left(d + q_{1-\frac{\alpha}{2}}\right) - \Phi\left(d - q_{1-\frac{\alpha}{2}}\right).$$

dove abbiamo posto

$$d = \frac{\mu_0 - \mu}{\sigma}\sqrt{n}.$$

La potenza del test è, per definizione, $1-\beta(\mu)$, cioè la probabilità di rifiutare $\mathcal{H}_0$ quando è falsa.

Viene poi calcolata la potenza del test unilaterale destro

$$\beta(\mu) := P_\mu (Z < q_{1-\alpha}) = ... = \Phi(d + q_{1-\alpha})$$

I due test, unilaterale e bilaterale, differiscono in potenza. Vedremo successivamente che quello unilaterale è più potente.

Lezione 20 (26/1). Vedere la registrazione di Barsanti. Regressione lineare: calcoli ed esempio. Tabelle di contingenza. Test per l'indipendenza. Coefficiente di determinazione.

Lezione 21 (29/1). Interpretazione grafica di β (e quindi della potenza), tramite il grafico della gaussiana canonica, il suo intervallo che corrisponde alla regione di accettazione, la gaussiana di media $-d$ (quella vera) e la probabilità che essa attribuisce alla regione di accettazione. Disegno nel caso bilaterale e unilaterale, da cui si intuisce che $\beta_{UD}(\mu) < \beta_B(\mu)$ se $\mu > \mu_0$. Si potrebbe in effetti dimostrare che la potenza unilaterale destra è maggiore della bilaterale quando $\mu > \mu_0$:

$$\Phi(d + q_{1-\alpha}) < \Phi\left(d + q_{1-\frac{\alpha}{2}}\right) - \Phi\left(d - q_{1-\frac{\alpha}{2}}\right)$$

(si noti che $\mu > \mu_0$ corrisponde a $d < 0$). La verifica analitica non è elementare, ma alcune cose si intuiscono. Ad esempio (i dettagli di queste spiegazioni non sono stati svolti a lezione), vale sempre $d + q_{1-\alpha} < d + q_{1-\frac{\alpha}{2}}$, quindi $\Phi(d + q_{1-\alpha}) < \Phi\left(d + q_{1-\frac{\alpha}{2}}\right)$. Allora, quando $\Phi\left(d - q_{1-\frac{\alpha}{2}}\right)$ è piccolissimo rispetto allo scarto tra $\Phi(d + q_{1-\alpha})$ e $\Phi\left(d + q_{1-\frac{\alpha}{2}}\right)$, è ragionevole che valga $\beta_{UD}(\mu) < \beta_B(\mu)$. Ora, siccome $d < 0$, $\Phi\left(d - q_{1-\frac{\alpha}{2}}\right)$ è sempre piuttosto piccolo, e diventa piccolissimo già per valori moderati di d, come $d = -2$ (la Φ è infinitesima per valori intorno a 3-4), se prendiamo come riferimento $\alpha = 0.05$, per esemplificare. Viceversa, per tali valori, $d + q_{1-\alpha}$ e $d + q_{1-\frac{\alpha}{2}}$ sono numeri abbastanza vicini a zero, quindi $\Phi(d + q_{1-\alpha})$ e $\Phi\left(d + q_{1-\frac{\alpha}{2}}\right)$ differiscono in modo non banale. Non sorprende quindi che per valori di d vicini a -2 la disuguaglianza $\beta_{UD}(\mu) < \beta_B(\mu)$ sia più netta. Altrove sfuma verso l'uguaglianza. Si veda il disegno, ottenuto nel caso $\alpha = 0.05$.

Curva caratteristica operativa: la curva della potenza (o di β) in funzione di $|d|$.

DOE (si veda l'analogo per la stima dei parametri nella lezione 15): in fase di progettazione degli esperimenti possiamo cercare di determinare la

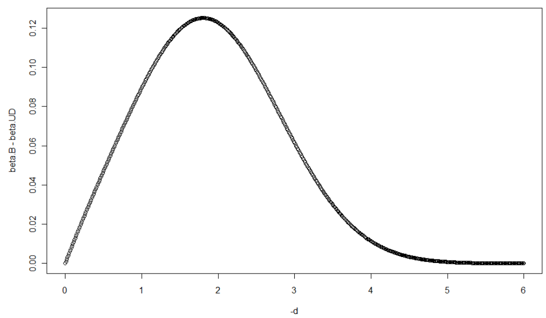

Figure 1: Grafico di $\beta_B - \beta_{UD}$ in funzione di -d, per $\alpha = 0.05$.

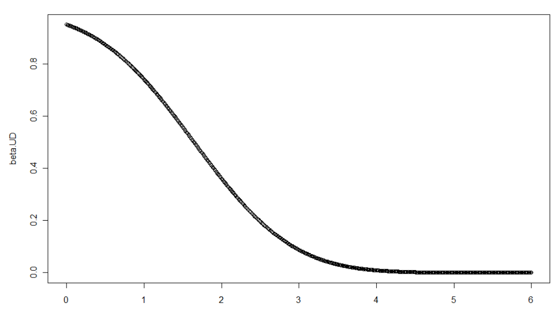

Figure 2: Grafico di β_{UD} in funzione di -d, per $\alpha = 0.05$.

numerosità n degli esperimenti che porterà ad avere una certa potenza, relativamente ad una significatività scelta ed a un valore di μ. Ad esempio, nel caso unilaterale destro, va risolta un'equazione del tipo $\Phi\left(d+q_{1-\alpha}\right)=0.1$ (se si vuole potenza 0.9), da cui $d+q_{1-\alpha}=q_{0.1}$, $d=-q_{0.9}-q_{1-\alpha}$, da cui si trova n ricordando che $d=\frac{\mu_0-\mu}{\sigma}\sqrt{n}$.

In una certa pratica ingegneristica vengono usate le curve OC proprio per questo. Le curve hanno il pregio di far ragionare con buon senso, circa il desiderio di miglioramento in rapporto al costo. Un ragionamento può essere: fino a circa $d=-3$ ogni aumento di $-d$ provoca un miglioramento netto della potenza, ma da lì in poi il tasso di miglioramento cala, serve un incremento sempre più ampio di $-d$ per ottenere piccoli miglioramenti della potenza. Anche se è un concetto, vago, è come se la curva avesse un "gomito". Allora accontentiamoci di $d=-3$. [Questo valore non è universale: è relativo alla curva OC per $\alpha=0.05$, e comunque è una nostra intuizione ad occhio.] Quindi $\frac{\mu_0-\mu}{\sigma}\sqrt{n}=-3$, da cui si trova n. La scelta $d=-3$ corrisponde alla potenza $1-\beta=1-\Phi\left(-3+1.64\right)=1-0.08=0.92$.

A livello pratico, ci sono molte scelte da fare, prima di calcolare n. Una delle più critiche è μ. Un'idea possibile, anche se vaga, è che μ sia il primo valore critico diverso da μ_0, cioè il primo valore che, se realizzato, provoca delle conseguenze rilevanti, e che quindi deve essere rilevato dal test. L'esempio seguente aiuta a capire questo concetto.

Carte di controllo. Quando si eseguono test nelle aziende? Ad esempio quando si fa monitoraggio, ad esempio con le *carte di controllo*. Vengono spiegate sommariamente le carte di controllo (vedere un cenno ai paragrafi 13.1-13.2), le bande, il campionamento a tempi regolari, l'allarme quando si esce dalle bande, l'analogia con l'eseguire un test ad ogni istante di controllo. Vengono calcolate le bande: $\mu_0\pm\frac{\sigma q_{1-\frac{\alpha}{2}}}{\sqrt{n}}$, come per l'intervallo di confidenza (ma lo scopo è un test, non la stima della media). Operativamente, pensando ad un esempio concreto, come si realizza una carta? Vanno scelti α ed n, fissati una volta per tutte. Per sceglierli si deve aver chiaro il significato di ogni elemento della teoria del test. α è la probabilità di uscire dalle bande quando invece la media è rimasta μ_0. In molti casi questo non è così grave: quando accade basta rifare il campionamento, per controllare meglio. n va scelto per avere una certa potenza, relativamente ad un certo μ. Si deve sapere, ad es. dagli esperti delle cose prodotte, quali deviazioni da μ_0 rendono inservibili o pericolose le cose prodotte. μ corrisponde a tali valori critici. A quel punto va scelto $\beta\left(\mu\right)$. E' la probabilità di non accorgersi di un cambiamento di μ_0, quando questo è avvenuto. Questo sì che può essere pericoloso: vendere cose fuori norma senza saperlo. Allora $\beta(\mu)$ va preso molto piccolo, e trovato n in corrispondenza.

Esempio: filato prodotto correttamente: $\mu_0 = 0.2$ mm, $\sigma_0 = 0.02$ mm. Spessore da evitare: 0.3 mm, o 0.1 mm. Creare carte di controllo.

Osservazione 1: si tratta di un processo ad alta precisione, cioè con σ_0 molto piccola. Un campione ha probabilità piccolissima di superare 0.3 mm per caso (servono 5 sigma, quasi impossibile).

Osservazione 2: ha quindi senso tenere sotto controllo la media, invece che il singolo esemplare. Infatti il singolo esemplare è improbabile che superi 0.3 mm per caso, se la media resta quella. Il pericolo non è nella causalità, ma in un peggioramento sistematico della media.

Osservazione 3: oppure il pericolo è in un peggioramento della varianza: se fosse ad es. $\sigma = 0.05$ mm., basta arrivare a 2σ per raggiungere 0.3 mm per caso in un esemplare. Questo sarebbe frequente.

Conclusione: vanno tenute sotto controllo media e varianza. Noi qui studiamo solo la media (chi vuole può vedere dei cenni sulle carte di controllo per σ nel capitolo 13).

Si crea una carta di controllo per la media, in cui $UCL = \mu_0 + \frac{\sigma_0 q_{1-\frac{\alpha}{2}}}{\sqrt{n}}$ e $LCL = \mu_0 - \frac{\sigma_0 q_{1-\frac{\alpha}{2}}}{\sqrt{n}}$. Vanno scelti α ed n. Come già detto sopra, α forse può essere scelto non troppo severo, es. 0.05. Tuttavia, immaginiamo di costruire un sistema di controllo automatico, che suona quando si superano le soglie: non vorremo che suoni per caso 5 volte su 100. Prendiamo allora $\alpha = 0.001$, ad esempio ($q_{1-\frac{\alpha}{2}} = 3.29$).

Invece n va scelto con lo scopo di avere una certa potenza. Va identificato un valore μ che non si vuole raggiungere. Va evitato che lo spessore sia 0.3 mm, o 0.1 mm. Trascurando l'influsso delle piccole fluttuazioni del singolo esemplare, va evitato che la media raggiunga questi valori. Quindi $\mu = 0.3$ oppure $\mu = 0.1$ sono i valori di riferimento rispetto a cui calcolare la potenza: essa misura la capacità della carta di controllo di accorgersi che lo spessore ha raggiunto quei livelli inaccettabili. Allora, scelta una potenza, es. 0.9999, cioè $\beta = 0.0001$, si impone l'equazione

$$\Phi(d + 3.29) - \Phi(d - 3.29) = 0.0001$$

dove $d = \frac{\mu_0 - \mu}{\sigma}\sqrt{n} = -\frac{0.1}{0.02}\sqrt{n} = -5\sqrt{n}$. Trascuriamo $\Phi(d - 3.29)$ che plausibilmente è molto piccolo, risolviamo $\Phi(d + 3.29) = 0.0001$, cioè $d + 3.29 = q_{0.0001} = -3.72$, $d = -3.29 - 3.72 = -7.01$, da cui $5\sqrt{n} = 7$, $\sqrt{n} = \frac{7}{5} = 1.4$,

$n = 2$. Questo risultato è un po' troppo incoraggiante rispetto a molti esempi applicativi, ed è dovuto al fatto che ci vogliono 5σ per passare causalmente da 0.2 a 0.3; cioè si tratta di un esempio con un grado di precisione già elevatissimo.

[Non svolto, ma segnalato per chi fosse interesato: poi si deve creare una carta di controllo per la varianza. Si può impostare così: l'indicatore è positivo, c'è un'unico limite, UCL. Un modo è prendere come indicatore $\frac{S^2}{\sigma_0^2}(n-1)$, e $UCL = \chi^2_{\alpha,n}$. Di nuovo vanno scelti α ed n, con ragionamenti analoghi.]

Complementi sulla prima parte del corso: mancanza di memoria dell'esponenziale, teorema sul minimo, entrambi dimostrati e commentati dal punto di vista dell'interesse applicativo. Si trovano entrambi nel paragrafo 5.6.