# Esercizi sulle metriche di prestazione

## 1 INTERVALLI DI CONFIDENZA

Una serie di **N** misure indipendenti ($x_1$, $x_2$, $x_3$, … $x_N$) di una determinata grandezza permette di stimare il valore medio **μ** di tale grandezza sull'intera popolazione. Siano allora **X** e $\boldsymbol{\sigma^2}$ rispettivamente il valore medio e la varianza delle misure.

L'intervallo ($\underline{X}$-ε , $\underline{X}$+ε) è detto intervallo di confidenza al *t* percento se *t* percento è la stima della probabilità che **μ** sia contenuto in esso. Naturalmente *t* ed ε sono fra loro correlati, al crescere della probabilità *t* l'ampiezza ε si riduce.

Diamo come noti i seguenti punti:

- La deviazione standard della media di N misure vale: $\sigma/\sqrt{N}$
- Al posto di σ si preferisce sostituire **s** = $\sigma \frac{N}{N-1}$ (**s** stima "unbiased" di σ, cioè tale che $E[s^2] = \sigma^2$)
- Al crescere di N la funzione $(x-\mu)\frac{\sqrt{N}}{s}$ tende in probabiità alla distribuzione normale di media 0 e varianza 1 denominata: *N*(0,1); vedi alcuni valori nella tabella:

| | | | | | | |
|---|---|---|---|---|---|---|
| 1-α | 0.9 | 0.95 | 0.975 | 0.99 | 0.995 | 0.999 |
| α | 0.1 | 0.05 | 0.025 | 0.01 | 0.005 | 0.001 |
| t(α) | 1.282 | 1.645 | 1.960 | 2.326 | 2.576 | 3.090 |

che è da interpretarsi come segue:

$\Pr\left\{(x-\mu)\frac{\sqrt{N}}{s} \le t(\alpha)\right\}$ = 1-α cioè: $\Pr\left\{x - t(\alpha)\frac{s}{\sqrt{N}} \le \mu\right\}$ = 1-α

A questo punto siamo in grado di affrontare il seguente problema:
Cinque misure indipendenti x(i), (i=1,2,.. 5) del tasso di arrivi delle transazioni in un certo ambiente forniscono:

| | | | | |
|---|---|---|---|---|
| 3.07 | 3.24 | 3.14 | 3.11 | 3.07 |

(N=5)

Possiamo dire che il tasso è > 3.00 con un livello di confidenza del 99.5 %? E se il livello scende al 97.5 %?

Allora: $\Sigma\, x(i) = 15.63$; $\Sigma\, x(i)^2 = 48.8791$; $\underline{x} = 3.126$; $\underline{x^2} = 9.77582$;

$\sigma^2 = \underline{x^2} - (\underline{x})^2 = 3.944 \times 10^{-3}$; $s^2 = \frac{5}{4}\sigma^2 = 4.93 \times 10^{-3}$; $s = 0.0702$

Fissato il livello di confidenza, dobbiamo verificare l'ipotesi che l'estremo inferiore del relativo intervallo sia > 3. Con i livelli del 99.5% e del 97.5% otteniamo rispettivamente:

$3.126 - \mathbf{2.576} \times 0.0702 / \sqrt{5} = 3.045$

$3.126 - \mathbf{1.960} \times 0.0702 / \sqrt{5} = 3.064$

Entrambi i valori sono maggiori di 3; come si è già detto, se abbassiamo il livello di confidenza l'estremo inferiore dell'intervallo si sposta verso l'alto.

Una soluzione più accurata si può ottenere se, al posto della distribuzione Normale, mettiamo la così detta *t di Student* per 4 gradi di libertà (4 = numero misure – 1), cinque misure sono infatti poche per usare la *N(0,1)*. Ecco allora la nuova tabella dove $t_4(\alpha)$ si riferisce alla *t di Student* con 4 gradi di libertà:

| | | | | | | |
|---|---|---|---|---|---|---|
| 1-α | 0.9 | 0.95 | 0.975 | 0.99 | 0.995 | 0.999 |
| α | 0.1 | 0.05 | 0.025 | 0.01 | 0.005 | 0.001 |
| $t_4(\alpha)$ | 1.533 | 2.132 | 2.776 | 3.747 | 4.604 | 7.173 |

Gli estremi inferiori dell'intervallo di confidenza divengono allora:

$3.126 - \mathbf{4.604} \times 0.0702 / \sqrt{5} = 2.981$

$3.126 - \mathbf{2.776} \times 0.0702 / \sqrt{5} = 3.039$

In conclusione se volessimo una confidenza del 99.5% dovremmo eseguire un maggior numero di misure ottenendo così una riduzione della varianza della media, quindi un'ampiezza minore per l'intervallo di confidenza oltre che un aumento dei gradi di libertà, che rende la distribuzione *t* ad avvicinarsi maggiormente alla *N(0,1)*.

# 2 TEST DELLE IPOTESI

Un argomento che si affronta in modo analogo al precedente è il confronto fra due valori medi stimati da due serie di misure indipendenti provenienti dalle popolazioni

$\{X\}$: x(1), x(2), ….x(n)    di media $\underline{x}$, stima di μ(X)

e

$\{Y\}$: y(1), y(2), ….y(m)    di media $\underline{y}$, stima di μ(Y)

L'algoritmo per determinare al livello α di significanza se accettare o rifiutare l'ipotesi μ(X) - μ(Y) = d contro l'alternativa:

a) μ(X) - μ(Y) > d
b) μ(X) - μ(Y) < d
c) μ(X) - μ(Y) ≠ d

richiede il calcolo della funzione

z = $\frac{(x-y)-d}{\sqrt{\frac{s_x^2}{n}+\frac{s_y^2}{m}}}$ che tende in probabilità alla *N(0,1)*

le regioni critiche sono nelle tre ipotesi precedenti:

a) z > t(α)
b) z < - t(α)
c) z < - t(α/2) ; t(α/2) < z

per esempio per α = 0.025 t(α) = 1.960 (valore che trovate in pag. 1)

Veniamo all'esempio pratico:

Una compagnia ha due sistemi che eseguono la stessa applicazione per due gruppi diversi di utenti, si vuole ottenere un bilanciamento dei tempi di risposta spostando eventualmente utenti da un sistema all'altro. Le misure che disponiamo sono le seguenti:

| sistema | numero misure | resp medio | std. dev |
|---|---|---|---|
| 1 | 100 | 20.24 | 5.6 |
| 2 | 120 | 18.72 | 4.2 |

Calcoliamo da queste il valore z:

z = $\frac{20.24-18.72}{\sqrt{\frac{5.6^2}{100}+\frac{4.2^2}{120}}}$ = 2.2397 > 2.326 → il sistema 1 ha un tempo di risposta maggiore del sistema 2 con un livello di confidenza del 99%

Se ora volessimo calcolare la distanza d fra i valori medi con un livello del 97.5%, cioè con α = 0.025, dobbiamo risolvere l'equazione:

1.960 = $\frac{20.24-18.72-d}{\sqrt{\frac{5.6^2}{100}+\frac{4.2^2}{120}}}$ ovvero d ≥ 0.2797 / $\sqrt{\frac{5.6^2}{100}+\frac{4.2^2}{120}}$ = 0.412