

Alcune considerazioni sulla misurazione delle spore

April 24, 2007

Va detto, in premessa, che la misurazione delle spore, può essere convenientemente validata, solo su spore da sporata, perchè sono le uniche mature.

Quando si ha a che fare con un campione, nel nostro caso di spore, dobbiamo avere ben presente che il campione “incorpora” almeno tre errori:

1) L'errore dovuto alla variabilità intrinseca del campione rispetto alla popolazione dal quale è estratto. Tale variabilità è relativa sia al fungo dal quale la sporata è estratta, sia alla popolazione della specie a cui il carpoforo appartiene

2) L'errore dell'osservatore, compiuto nel momento in cui misura

3) L'errore dello strumento (microscopio) utilizzato per le rilevazioni

Per quanto l'osservatore si sforzi di mettere accuratezza nella propria ricerca (accuratezza dovuta ed auspicabile) questi tre tipi di errore sono **ineliminabili**.

Dunque, gli errori, se non eliminabili, devono essere attentamente “pesati”

Quasi tutti i testi di micologia, in riferimento alla misurazione delle spore, riportano una scrittura di questo tipo:

$$Lunghezza = (8);9\div 11,5;(13)\mu m(1)$$

Con tale notazione si evidenzia che è stata rinvenuta una lunghezza minima di (8); un intervallo in cui ricadono la maggior parte delle lunghezze, pari a $9\div 11$ e una misura massima pari a (13) μm .

Analoga scrittura è riservata alla larghezza.

Se da un punto di vista pratico le differenze con altri metodi di notazione non sono apprezzabili, lo sono dal punto di vista teorico; per due motivi:

a) La notazione in questione è relativa a una misurazione campionaria: descrive cioè semplicemente alcuni valori rilevati in un generico campione c_k senza che siano state compiute inferenze sulla popolazione

b) E' privo di significato quantificare in modo separato “lunghezza” e “larghezza” delle spore; in quanto esse vanno rilevate (ed analizzate) in modo congiunto.

In pratica dobbiamo ricorrere alla statistica bivariata e rilevare coppie dei caratteri “lunghezza” (X) e “larghezza” (Y) ; vale a dire:

$$x_1y_1; x_2y_2....; x_ky_k....; x_ny_n$$

Trascuriamo per un attimo la “larghezza” e supponiamo di volere rilevare solo le misure della “lunghezza”.

Immaginiamo di avere a disposizione la sporata di un certo carpoforo e che quattro campioni vengano da essa estratti e consegnati ad altrettanti osservatori.

Supponiamo anche che di ogni campione si vogliano osservare 30 misure (N =30) . A lavoro terminato avremo la seguente situazione:

$$C_1 \text{con} \overline{X}_1 S_1$$

$$C_2 \text{con} \overline{X}_2 S_2$$

$$C_3 \text{con} \overline{X}_3 S_3$$

$$C_4 \text{con} \overline{X}_4 S_4$$

Dove: $\overline{X}$ è la media campionaria e S è lo scarto quadratico medio (o deviazione standard) del campione
E' noto che la media aritmetica è definita come:

$$\overline{X} = \frac{\sum x_i}{n}$$

e lo scarto quadratico medio come:

$$S = \sqrt{\frac{(x_i - \bar{x})^2}{n-1}}$$

In ragione degli errori presenti nella rilevazione (con le caratteristiche evidenziate in premessa) molto probabilmente si avrà una situazione del tipo:

$$C_1 \neq C_2 \neq C_3 \neq C_4$$

Senza gli errori (cioè una generica variabile casuale in grado di riassumere tutte le accidentali fonti di variabilità) I campioni sarebbero uguali, in quanto estratti dalla medesima popolazione per cui

$$C_1 - C_2 = 0; C_1 - C_3 = 0 \text{ ecc.}$$

Ne consegue che ogni campione, avendo media e varianza diversa da ogni altro, non può essere, di per se, rappresentativo della popolazione da cui è estratto.

Se non lo è media e varianza, non lo sono , a maggior ragione, i valori minimo e massimo posti all'estremo , rispettivamente, sinistro e destro della notazione (1)

giacchè siamo autorizzati a ritenere che ogni campione casuale possa avere valori minimi e massimi diversi da qualsiasi altro campione casuale, estratto dalla medesima popolazione.

Va notato che noi non siamo interessati ad ottenere dati campionari fini a se stessi; ma dati campionari in quanto rappresentativi dell'intera popolazione.

Ne consegue che, noti i dati campionari, è necessario compiere inferenze sulla popolazione costruendo l'intervallo della media vera.

Essendo incognita la varianza σ^2 della popolazione, dovremo ricorrere al "t di Student" la cui formula risolutiva è:

$$\mu = \bar{X} \pm t_{\frac{\alpha}{2}(n-1)} \cdot \frac{S}{\sqrt{n}}$$

dove:

- n è il numero di dati

- S è la deviazione standard calcolata sui dati del campione

- $t_{\frac{\alpha}{2}(n-1)}$ è il valore tabulato di t ad un prestabilito rischio α con n-1 gradi di libertà.

Esempio:

sia rilevata da un osservatore la lunghezza di 14 spore di Xerocomus crysenteron, come riportato in tabella:

Table 1: Spore di Xerocomus crysenteron

	Unità stat.	Valore osserv.	Media	Scarto	$Scarto^2$
			$\bar{X} = \frac{\sum x_i}{n}$	$X_i - \bar{X}$	$(X_i - \bar{X})^2$
	1	14,25	13,18	1,07	1,15
	2	13,04	13,18	-0,14	0,02
	3	12,36	13,18	-0,82	0,67
	4	14,45	13,18	1,27	1,62
	5	14,23	13,18	1,05	1,11
	6	13,83	13,18	0,65	0,43
	7	11,96	13,18	-1,22	1,48
	8	11,94	13,18	-1,24	1,53
	9	13,25	13,18	0,07	0,01
	10	12,34	13,18	-0,84	0,7
	11	14,42	13,18	1,24	1,55
	12	11,90	13,18	-1,28	1,63
	13	13,09	13,18	-0,09	0,01
	14	13,40	13,18	1,22	0,05
Media	$\bar{X} = \frac{\sum x_i}{n}$	13,18			
Devianza	$\sum (x_i - \bar{X})^2$				11,94
Dev. Standard	$S = \sqrt{\frac{\sum (x_i - \bar{X})^2}{n-1}}$				0,27
Varianza	S^2				0,07

Sostituendo i dati alla (1) si avrà:

$$\mu = \bar{X} \pm t_{\frac{\alpha}{2}(n-1)} \cdot \frac{S}{\sqrt{n}}$$

sulle tavole statistiche $t_{0,025}13g.d.l = 2,53$
sostituendo:

$$\mu = 13,18 \pm 2,53 \cdot \frac{0,27}{\sqrt{14}} = 13,00 \leq \mu \leq 13,36$$

Ciò significa che il valore vero della lunghezza è compreso nell'intervallo tra

$$13,00 \div 13,36 \mu m$$

E' facile dimostrare che, per $N \geq 10$ unità statistiche il t di Student perda di significato e che, in sostituzione dello stesso, si possa usare il valore "Z" al rispettivo rischio α della distribuzione normale standardizzata N (0,1)

L'intervallo della media vera (della popolazione) con formulazione di rischio $1-\alpha$, con $\alpha = 0,05$ (prob. 95%: valore tabulato Z= 1,96)

$$13,18 - 1,96 \cdot \frac{0,27}{\sqrt{14}} \leq \mu \leq 13,18 + 1,96 \cdot \frac{0,27}{\sqrt{14}}$$

ovvero:

$$13,03 \leq \mu \leq 13,32$$

Valori molto prossimi a quelli individuati col t di Student.
Dunque la scrittura completa e corretta sarà:

Table 2: Dati caratteristici della lunghezza delle spore di *Xerocomus crysenteron*

Indicatori	Valori
Numero di osservazioni	14
Media campionaria	13,18
Dev. std. del campione	0,27
Rischio α	0,05 (95% prob.)
Intervallo di confidenza	$13,03 \leq \mu \leq 13,32$

Analoga procedura deve essere compiuta per la larghezza.

Successivamente, ma questo è un'altro aspetto del problema è bene considerare come caratteri congiunti (xy) la lunghezza e la larghezza delle spore; valutare il loro grado di associazione e gli altri parametri della regressione, tramite gli strumenti della statistica bivariata.