

## Come si valuta la “normalità” di una distribuzione di dati?

- Come possiamo giudicare se i dati provengono da una distribuzione che può essere approssimata con una normale?
- Gli istogrammi, i diagrammi ramo-foglia e alcuni indici possono rivelare caratteristiche tipicamente non normali:  
outlier, asimmetria, interruzioni dei valori (gap), clusters.
- Se i grafici appaiono abbastanza simmetrici e unimodali occorre un metodo più sensibile, che possa rivelare l'adeguatezza del modello normale (simmetria, outlier, **peso delle code**).



## Plot dei quantili normali. Come si costruisce

Vogliamo verificare se un determinato campione proviene da una distribuzione normale (con ugual media e scarto st.)

1. Si ordinano le osservazioni , e si calcolano i **percentili campionari  $x_i$** .
2. Si considera la distribuzione normale standard e si trovano **i valori  $z_i$  che corrispondono agli stessi percentili (quantili normali standardizzati)**
3. Si costruisce un diagramma di dispersione con le osservazioni  $x_i$  sull'asse orizzontale e i corrispondenti quantili normali standardizzati  $z_i$  sull'asse verticale
4. Si verifica la normalità delle osservazioni controllando se i punti del diagramma si trovano approssimativamente su una retta

Tutti i software statistici riportano i plot dei quantili normali (normal plots).

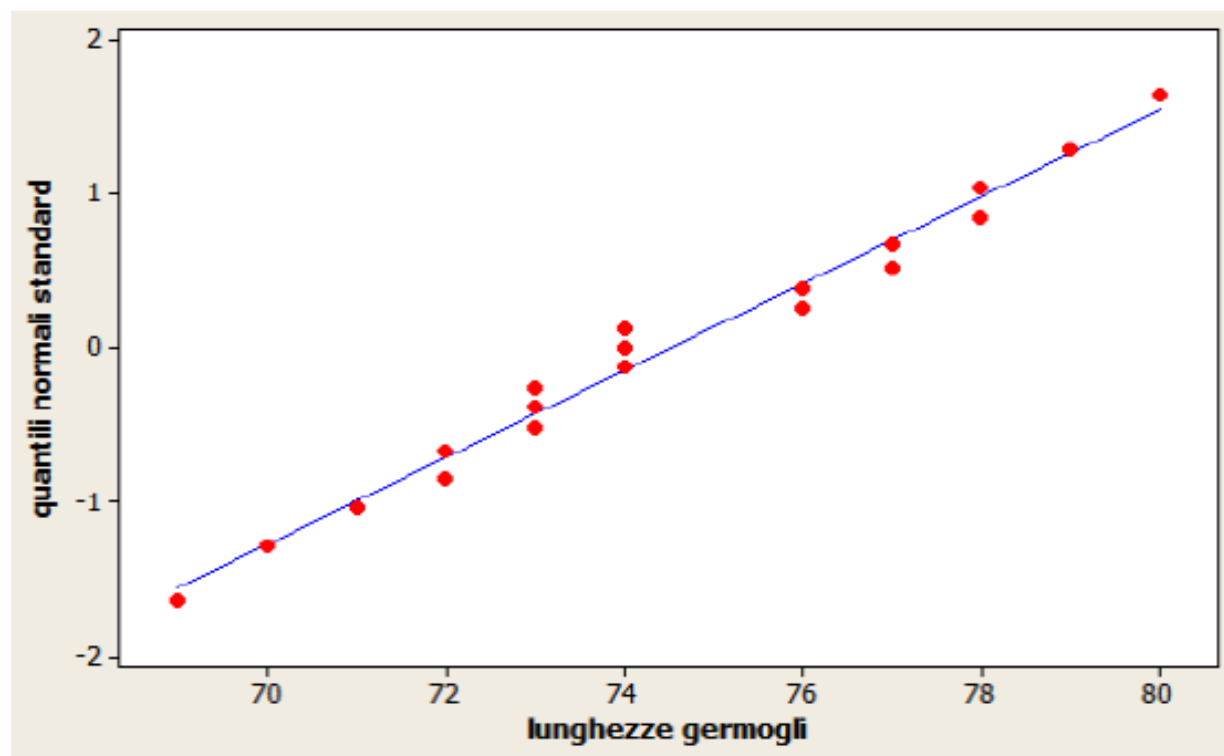


Plot dei quantili normali: metodo grafico di controllo della normalità dei dati

	$x_i$	$z_i$	
	69	-1,64485	
	70	-1,28155	
	71	-1,03643	
	72	-0,84162	
26-mo → percentile	72	-0,67449	← (quinto quantile) 26-mo percentile
	73	-0,52440	
	73	-0,38532	
	73	-0,25335	
lunghezze germogli	74	-0,12566	← quantili normali standardizzati
→	74	0,00000	←
	74	0,12566	
	76	0,25335	
	76	0,38532	
	77	0,52440	
	77	0,67449	
	78	0,84162	
	78	1,03643	
	79	1,28155	
	80	1,64485	← (diciannovesimo quantile) 95-mo percentile



## Plot dei quantili normali per l'es. germogli: diagramma di dispersione



I punti del normal plot si dispongono approssimativamente su una retta inclinata positivamente. Le lunghezze dei 19 germogli hanno una distribuzione approssimativamente normale.

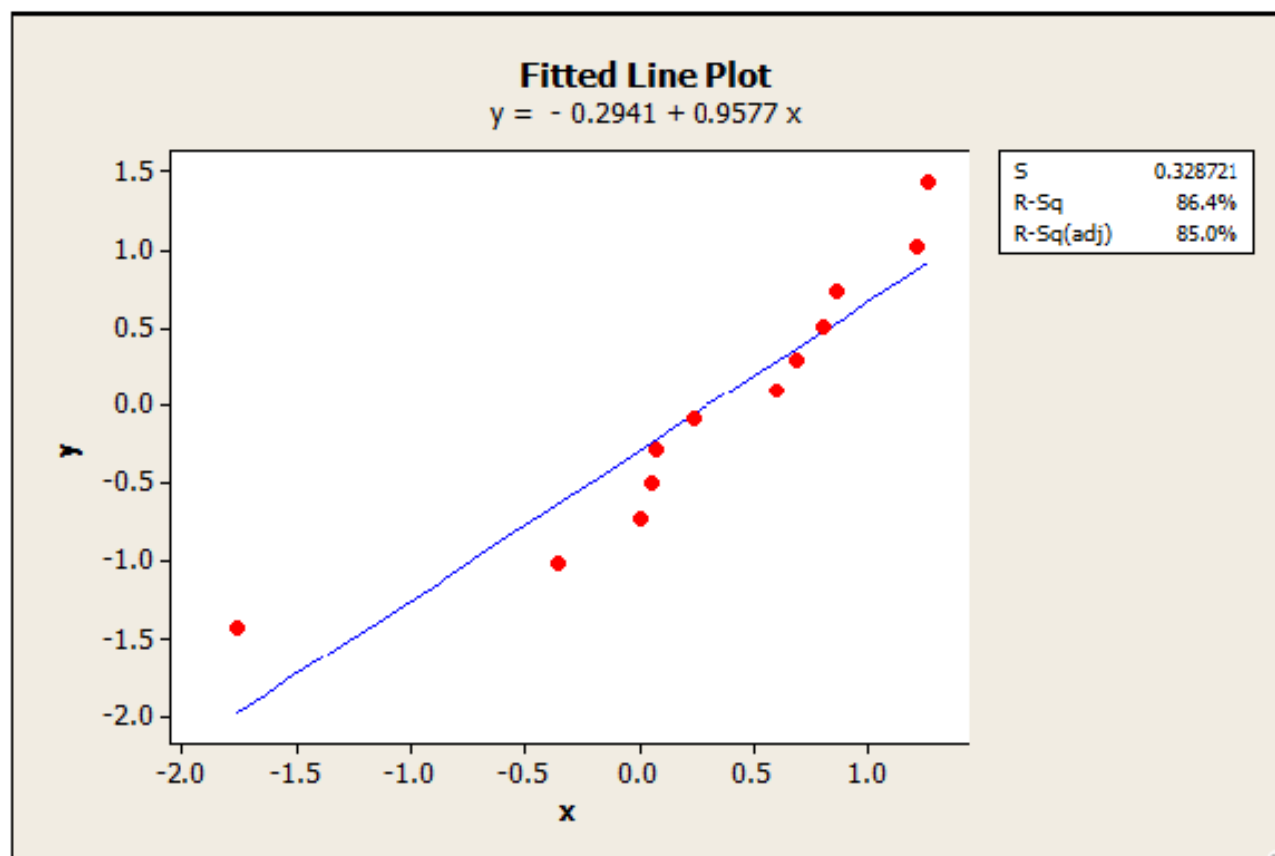


Plot dei quantili normali. Es.: campione di 13 dati estratto da una distribuzione normale standard

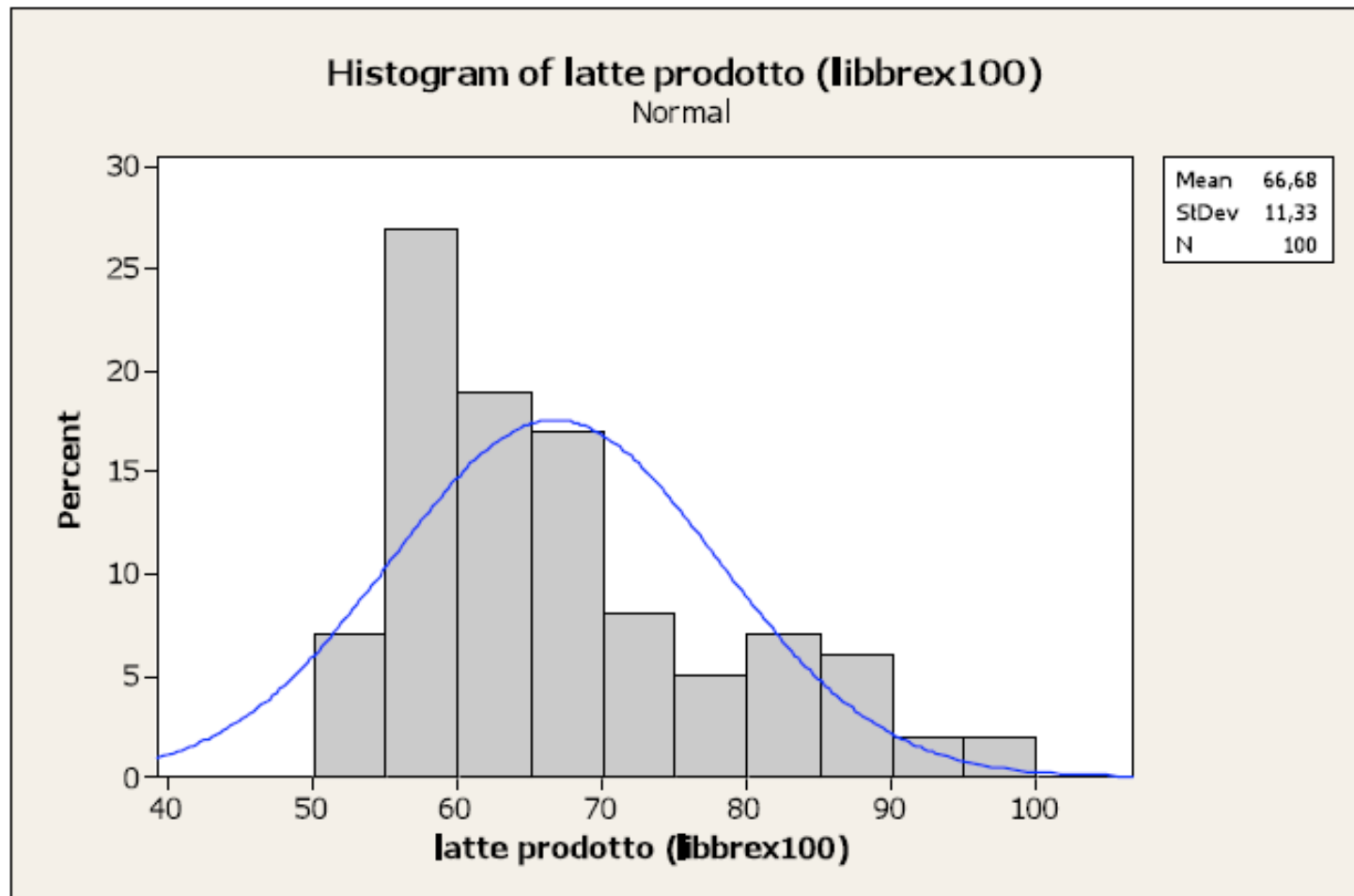
<b>x</b>	<b>percentili</b>	<b>y=z</b>
-1.75761	0.0769	-1.42624
-0.35848	0.1538	-1.02027
-0.00063	0.2308	-0.73621
0.04745	0.3077	-0.50238
0.06633	0.3846	-0.29342
0.23864	0.4615	-0.09666
0.60580	0.5385	0.09666
0.69193	0.6154	0.29342
0.81182	0.6923	0.50238
0.86228	0.7692	0.73621
1.21230	0.8462	1.02027
1.26512	0.9231	1.42624
1.42839	1.0000	*



Plot dei quantili normali. Es.: campione di 13 dati estratto da una distribuzione normale standard

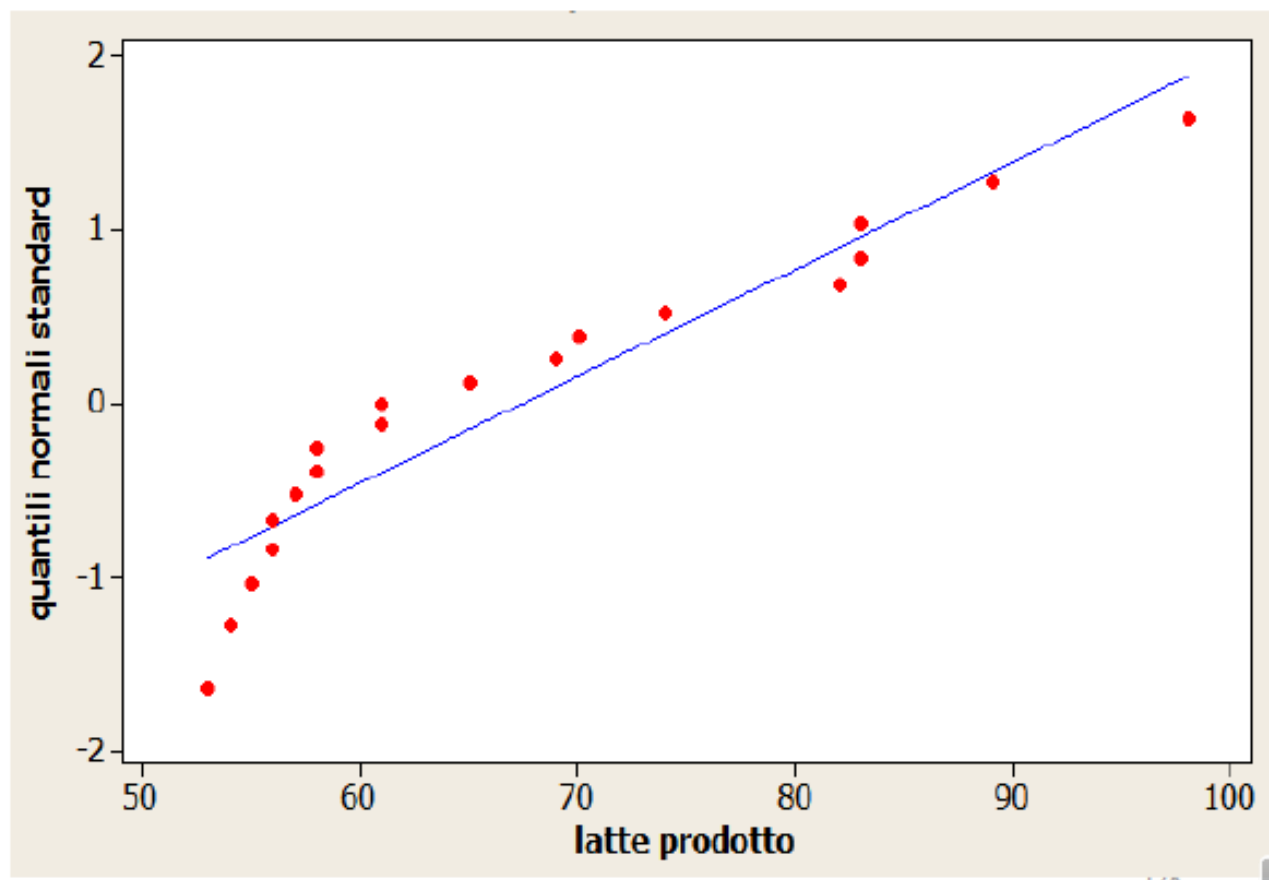


Produzione annuale di latte (libbre x 100) di 100 vacche Jersey di 2 anni (Sokal, Rohlf p.104)



**Plot dei quantili normali**  
**Diagramma di dispersione per 19 valori estratti dai 100**

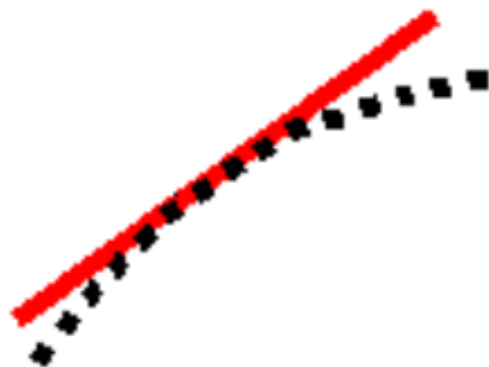
- 53
- 54
- 55
- 56
- 56
- 57
- 58
- 58
- 61
- 61
- 65
- 69
- 70
- 74
- 82
- 83
- 83
- 89
- 98







I dati presentano una  
lunga coda a sinistra

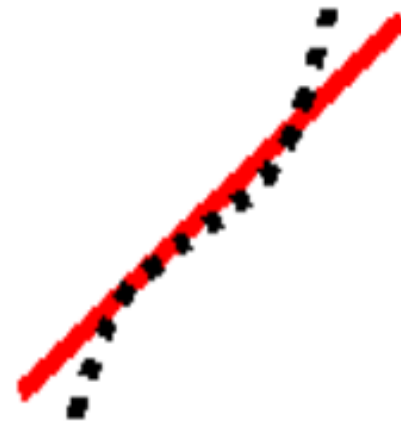


I dati presentano una  
lunga coda a destra





Code corte rispetto alla normale, ossia varianza minore rispetto a una normale



Code lunghe rispetto alla normale, ossia varianza maggiore rispetto a una normale

