Bootstrap statistical approach using R software to evaluate multimodal quantitative results in food microbiology PTs
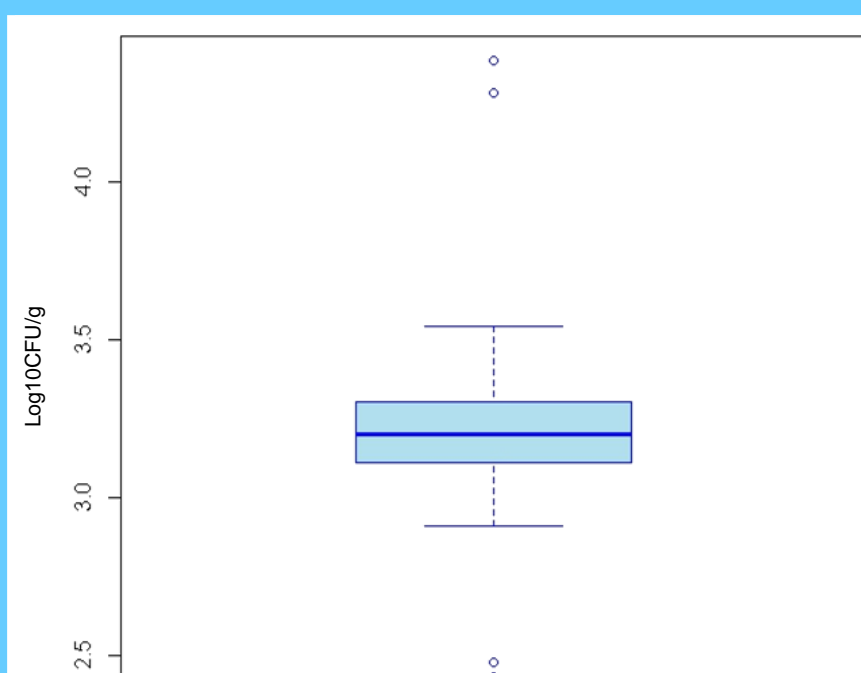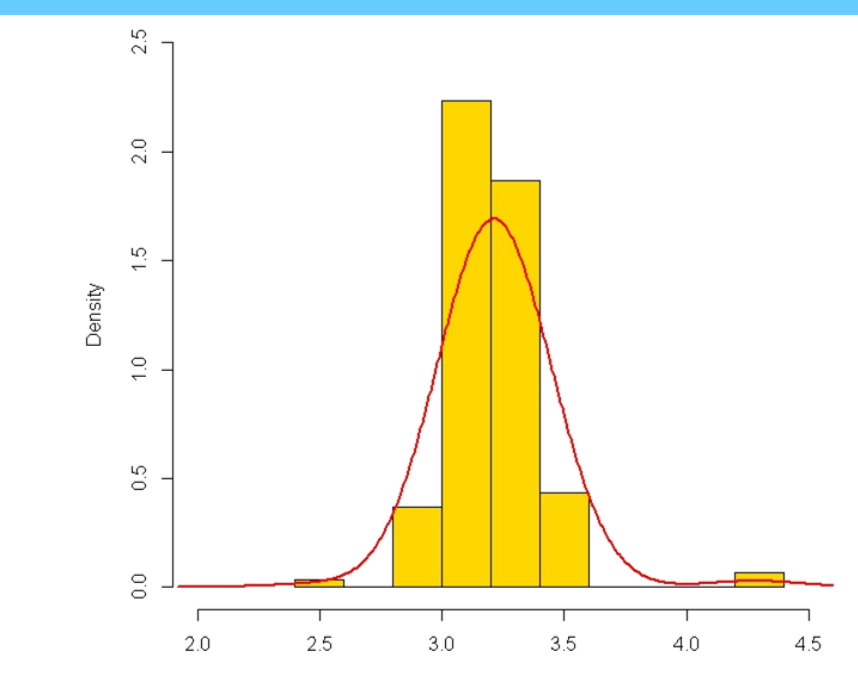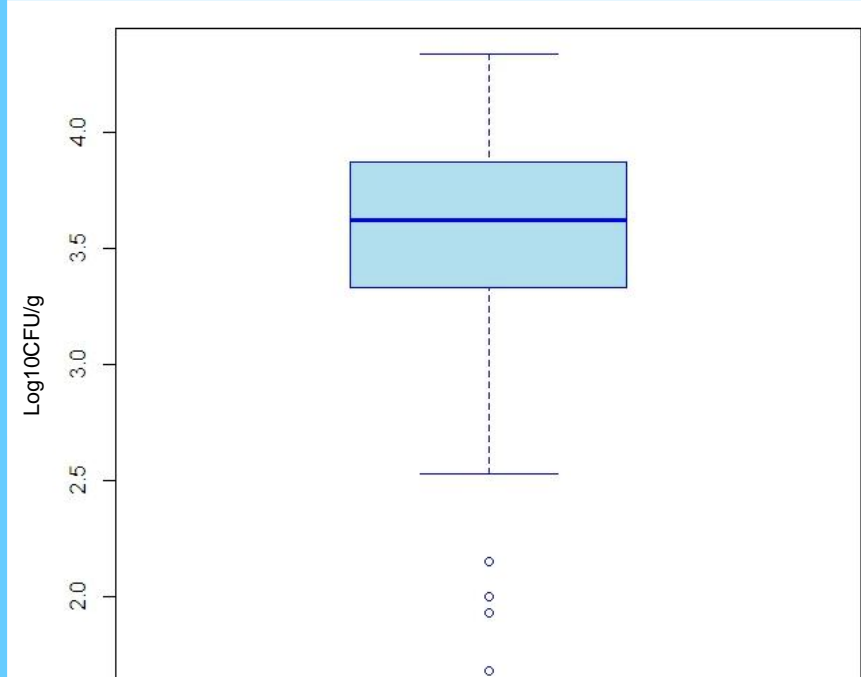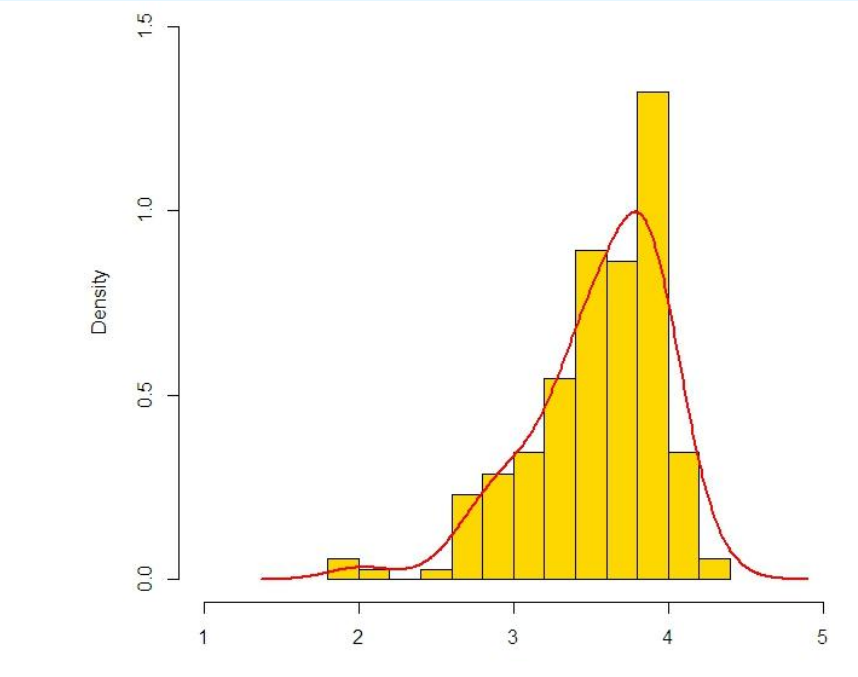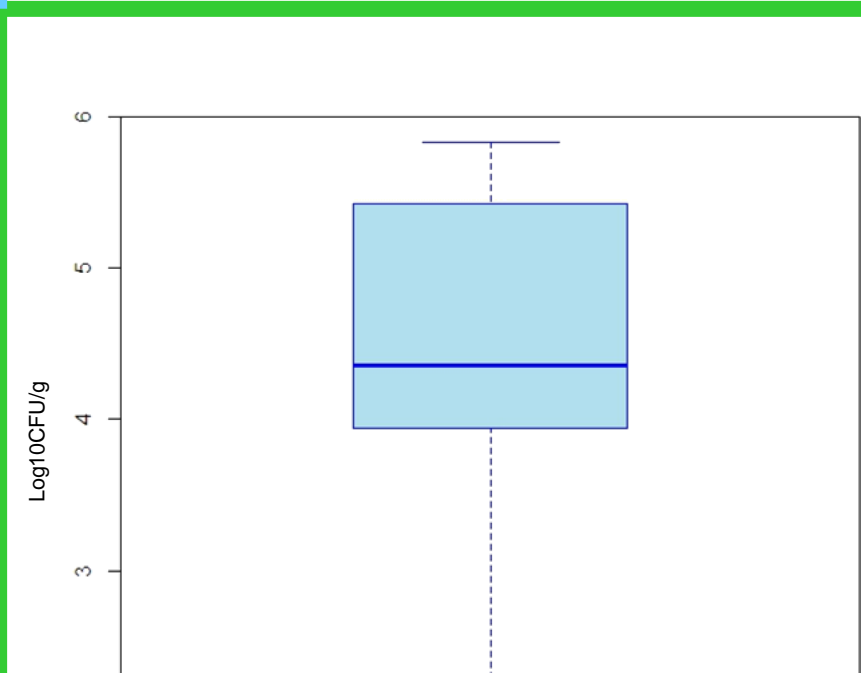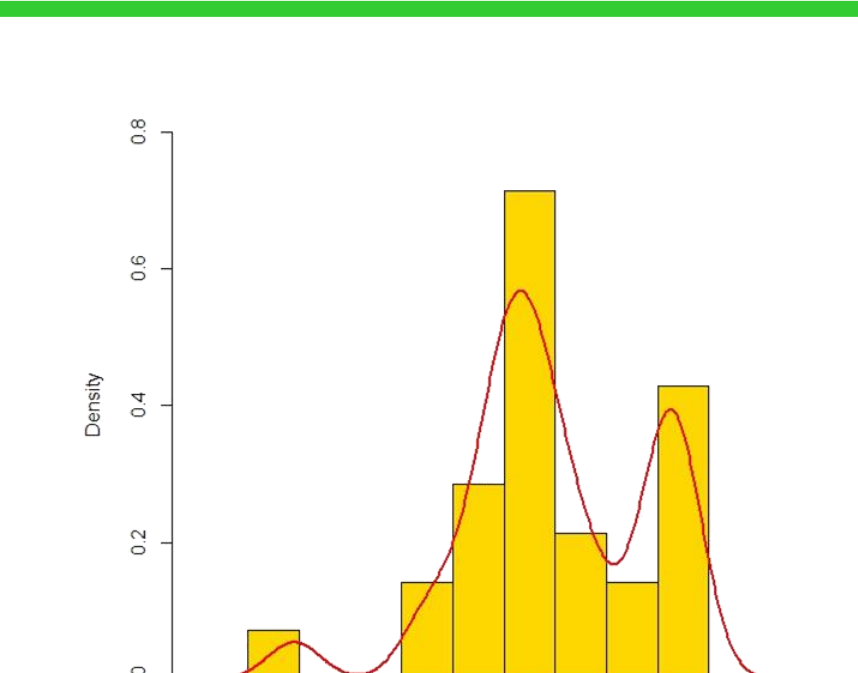
Marzia Mancin, Marica Toson, Maria Grimaldi, Lisa Barco, Romina Trevisan, Paola Carnieletto, Antonia Ricci

Istituto Zooprofilattico Sperimentale delle Venezie, Viale dell'Università 10, 35020, Legnaro (Padua), Italy

Statistical analysis has an essential role in the PT evaluation especially when the distribution of quantitative results is multimodal or strongly asymmetric, outliers aside. A possible solution is to estimate the modes of kernel density function of data distribution by using the bootstrap technique. The case of *Clostrifdium perfringens* (CFU/g) PT, accredited according to ISO/IEC 17043:2010 by the Italian Accreditation Body "Accredia".

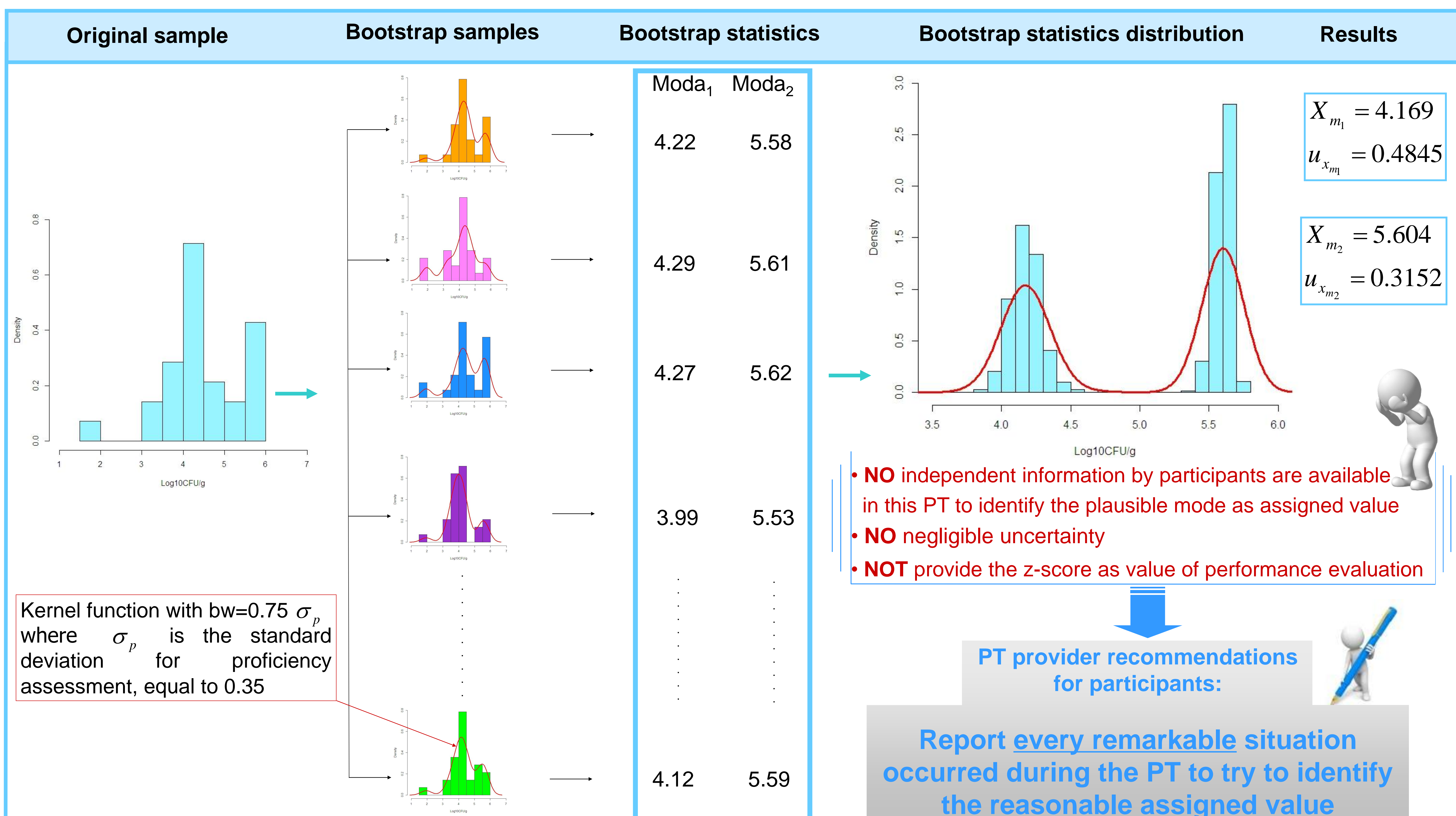## How was the data distribution of last *Clostridium perfringens* PT?

Box plot	Histogram and kernel function density	Conditions without outliers	Procedure
		Unimodal and simmetric or slightly asimmetric	• X^* = Robust average by Algorithm A of ISO 13528 • σ_p for proficiency assessment • $u_{x^*} = \sigma_p / \sqrt{n}$ • Uncertainty negligible condition $u_{x^*} < 0.3 \cdot \sigma_p$ • z-score to performance evaluation
		Unimodal but with a very strong asimmetry	• X_m = bootstrap mode • σ_p for proficiency assessment • $u_{x_m} = se_{x_m}$ • Uncertainty negligible condition $u_{x_m} < 0.3 \cdot \sigma_p$ • z-score to performance evaluation
		Multimodal	• X_{m_1} and X_{m_2}: bootstrap modes • σ_p for proficiency assessment • $u_{x_{m_1}} = se_{x_{m_1}}$ and $u_{x_{m_1}} = se_{x_{m_2}}$ • z-score to performance evaluation if <u>the plausible mode is identificable from available independent information</u> by participants and its uncertainty is negligible $u_{x_{m_i}} < 0.3 \cdot \sigma_p$

box plot or histograms or dot plots often do not show the true data distribution:
kernel density plot RECOMMENDED!!!

Analysis of *Clostridium perfringens* multimodal data

Original sample	Bootstrap samples	Bootstrap statistics	Bootstrap statistics distribution	Results
		Moda₁ Moda₂ 4.22 5.58 4.29 5.61 4.27 5.62 3.99 5.53 ⋮ ⋮ 4.12 5.59		$X_{m_1} = 4.169$ $u_{x_{m_1}} = 0.4845$ $X_{m_2} = 5.604$ $u_{x_{m_2}} = 0.3152$

Kernel function with bw=0.75 σ_p where σ_p is the standard deviation for proficiency assessment, equal to 0.35

• **NO** independent information by participants are available in this PT to identify the plausible mode as assigned value
• **NO** negligible uncertainty
• **NOT** provide the z-score as value of performance evaluation

PT provider recommendations for participants:

Report <u>every remarkable</u> situation occurred during the PT to try to identify the reasonable assigned value

Bibliography:
• Thompson M., Ellison S.L.R., Wood R., "*The International Harmonised Protocol for the Proficiency Testing of Analytical Chemistry Laboratories (IUPAC Technical report)*"
• ISO 13528 Statistical methods for use in proficiency testing interlaboratory comparisons