



KAREN L. SMITH, MD, MPH
Director & State Health Officer

State of California—Health and Human Services Agency
California Department of Public Health



EDMUND G. BROWN JR.
Governor

January 5, 2016

TO: Participants in the November 2015 Proficiency Test in Forensic Alcohol Analysis

SUBJECT: Assigned Values and Acceptable Ranges of Results for the November 2015 Proficiency Test in Forensic Alcohol Analysis

Attached is a summary of the descriptive statistics for the November 2015 proficiency test in forensic alcohol analysis. Included here are the target formulation values, the true values as determined by the Department's analyses, the peer-group or consensus values and the standard deviations, and graphical summaries of the distribution of participant results.

Historically, the Department has determined the acceptable limits of performance based on reported results that are within the range representing $\pm 5\%$ of the 99% confidence interval of the peer group mean, where the range has been truncated to two significant figures (Table 1). This range is described as the "Tier #2 interval." The Department also calculates a "Tier #1 interval," which represents the range of reported results that are within $\pm 5\%$ of the 95% confidence interval of the peer group mean where the range is based on the results reported to three significant figures. Tier #1 is expected to include those laboratories demonstrating a high degree of accuracy. The second, wider tier would include those laboratories not as close to the central tendency as the first tier, but still accurate and therefore adequately competent. Again, historically, the Department has used the wider second tier to evaluate the laboratories' results.

The IUPAC International Harmonized Protocol for the Proficiency Testing of Analytical Chemistry Laboratories (Harmonized Protocol) recommends the use of z-scores for evaluating proficiency test data. However, the Harmonized Protocol notes that the interpretation of the z-scores is based on the normal distribution of reported results, in which case the z-scores can be expected to follow the standard normal distribution. As indicated in Table 2, the results for Pools 10195 and 10265 in this proficiency test were not found to be normally distributed. Accordingly, the use of z-scores may not be completely appropriate, but they still may be useful to identify outlier and/or warning level results. The expression for calculating a z-score is included in Table 2. Generally a score between -2 and +2 ($|z| \leq 2$) is considered satisfactory or acceptable. A score outside the range -3 to +3, inclusive ($|z| \geq 3$) is considered unsatisfactory or unacceptable and the laboratory must take corrective actions. Z-scores between -3 and -2 or +2 and +3 ($2 < |z| < 3$) are considered questionable and these two ranges should be used as warning limits. Scores within the warning limit ranges in two or more consecutive test events could be considered unacceptable.

The proficiency test results expressed as z-scores for the participants whose results were used to determine the peer group mean and statistics in the March 2015 test are summarized in Figure 4. Participants are identified by codes. An attachment to this letter provides codes for participants from your laboratory. The figure is provided for educational purposes only and was not used to formally evaluate a laboratory's performance.

Another approach for evaluating proficiency test data, which is non-parametric and does not require the data to be converted to a standard normal form, divides the test data at regular intervals or quantiles¹. The quartile is a type of quantile: the first quartile (Q_1) is defined as the middle number between the lowest number and the median of the data set. The second quartile (Q_2) is the median of the data set. The third quartile (Q_3) is the middle value between the median and the highest value of the data set. The interquartile range (IQR), a measure of the dispersion of the data, is the difference between the upper and lower quartiles ($IQR = Q_3 - Q_1$). Boundaries (called fences) are set at $Q_1 - 1.5 IQR$ (lower fence) and $Q_3 + 1.5 IQR$ (upper fence) to identify potential outliers in the tails of the distribution. In Figure 3, the data from the two pools are presented as box and whisker or Tukey plots with the quartiles and fences shown. The median of the data is shown by a black line and the mean of the data is shown by a red line inside the box. These figures can be used by the participants to evaluate their data.

Sincerely,

Original signed by

Clay Larson, Chief
Abused Substances Analysis Section
Food and Drug Laboratory Branch

¹ Statistics and Chemometrics for Analytical Chemistry Sixth Edition, Miller and Miller (p. 158)

Statistical Data for November 2015 Proficiency Test in Forensic Alcohol Analysis

Table 1 **CDPH Tier #1 and Tier #2 Acceptable Ranges**

<u>Pool</u>	<u>Peer Group Mean</u>	<u>Tier #1</u>	<u>Tier #2</u>
#1	0.137	0.128 – 0.146	0.12 – 0.14
#2	0.226	0.212 – 0.240	0.21 – 0.24

Table 2 **Summary of Test Pool Data**

Parameter	POOL 1 (10195)		POOL 2 (10265)	
Pre-distribution Data	Target Value	0.14%	Target Value	0.23%
	True Value ²	0.138	True Value ¹	0.228
	Standard Deviation ¹	0.0007	Standard Deviation ¹	0.0017
Descriptive statistics	Mean	0.136	Mean	0.226
	Adjusted Mean ³	0.137	Adjusted Mean ²	0.226
	Standard Error ⁴	0.0006	Standard Error ³	0.0005
	Median	0.137	Median	0.226
	Standard Deviation	0.0040	Standard Deviation	0.0036
	Minimum	0.122	Minimum	0.214
	Maximum	0.143	Maximum	0.236
	Count	52	Count	52
Descriptive statistics (box plot)	Q1 (25%)	0.135	Q1 (25%)	0.224
	Q3 (75%)	0.138	Q3 (75%)	0.228
	IQR	0.0045	IQR	0.004
	Lower Fence	0.1305	Lower Fence	0.218
	Upper Fence	0.1425	Upper Fence	0.234
Histogram	Figure 1		Figure 2	
Normal distribution? ⁵	No (p<0.001)		No (p=0.007)	
Box Plot (SigmaPlot)	Figure 3		Figure 3	
Robust mean, X^{*6}	0.137		0.226	
Robust standard deviation, σ_{rob}	0.0018		0.0027	
Fitness-for-purpose standard deviation, σ_p^7	0.0034		0.0056	
Consensus value (X_a) determined as Mode ($\mu_{1/2}$) of Gaussian Kernel distribution	0.1372		0.2259	
Uncertainty of the consensus value, X_a , S.E. ⁸	0.0003		0.0004	
$X_a \pm$ S.E.	0.1372 ± 0.0003		0.2259 ± 0.0004	
z-score	$z = \frac{X - X_a}{\sigma_p}$		$z = \frac{X - X_a}{\sigma_p}$	

² Based on CDPH's Headspace Gas Chromatographic Method

³ Mean determined from participant data after the removal of outlier(s)

⁴ Standard Error of the Mean

⁵ Shapiro-Wilk test used at 0.05 significance level.

⁶ Robust average of the results reported by the participants was calculated using Algorithm A in Annex C of ISO 13528:2005.

⁷ The Department has determined a value for σ_p of 2.5% based on the uncertainties associated with the reported results on recent tests together with the 5% accuracy and precision standard of performance requirements set forth in the regulations.

⁸ Determined as Standard Error of Mode using bootstrap simulation technique with bandwidth of $0.75 \cdot \sigma_p$

Figure 1

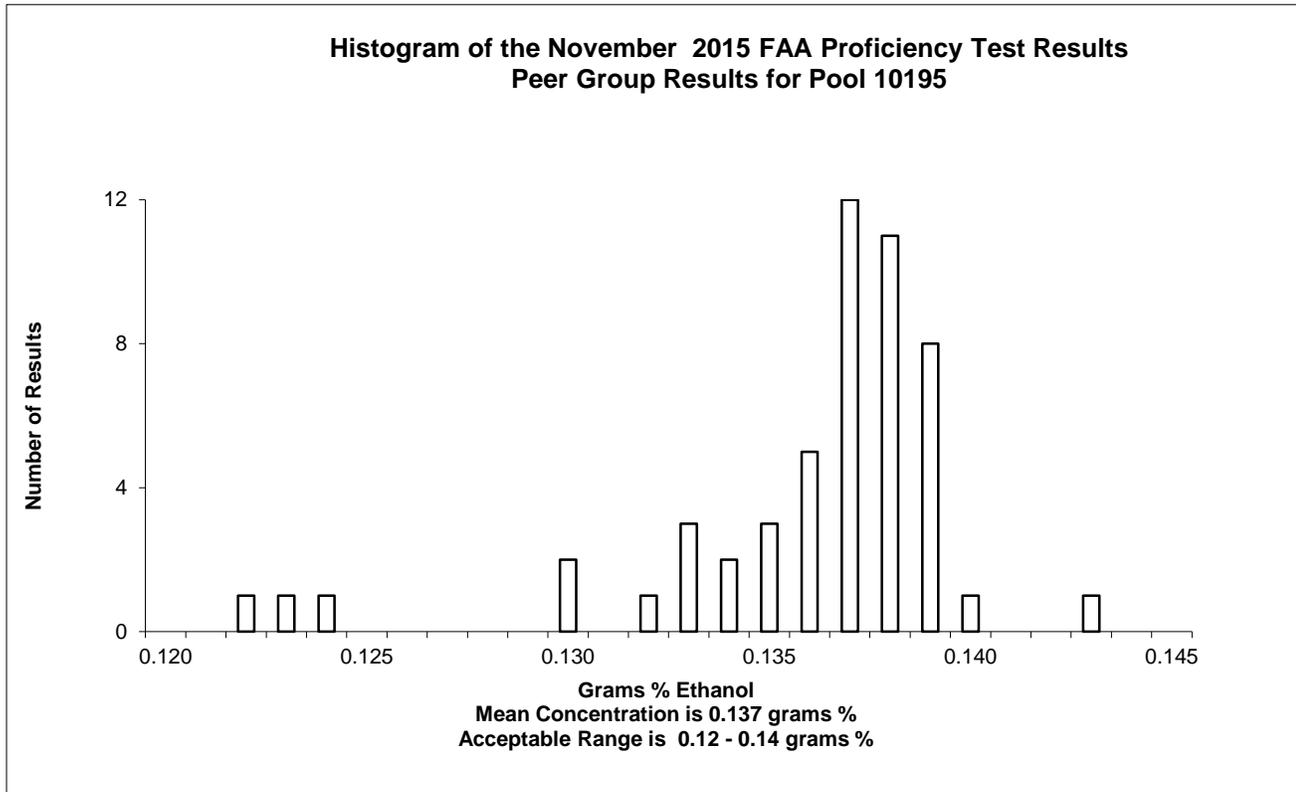


Figure 2

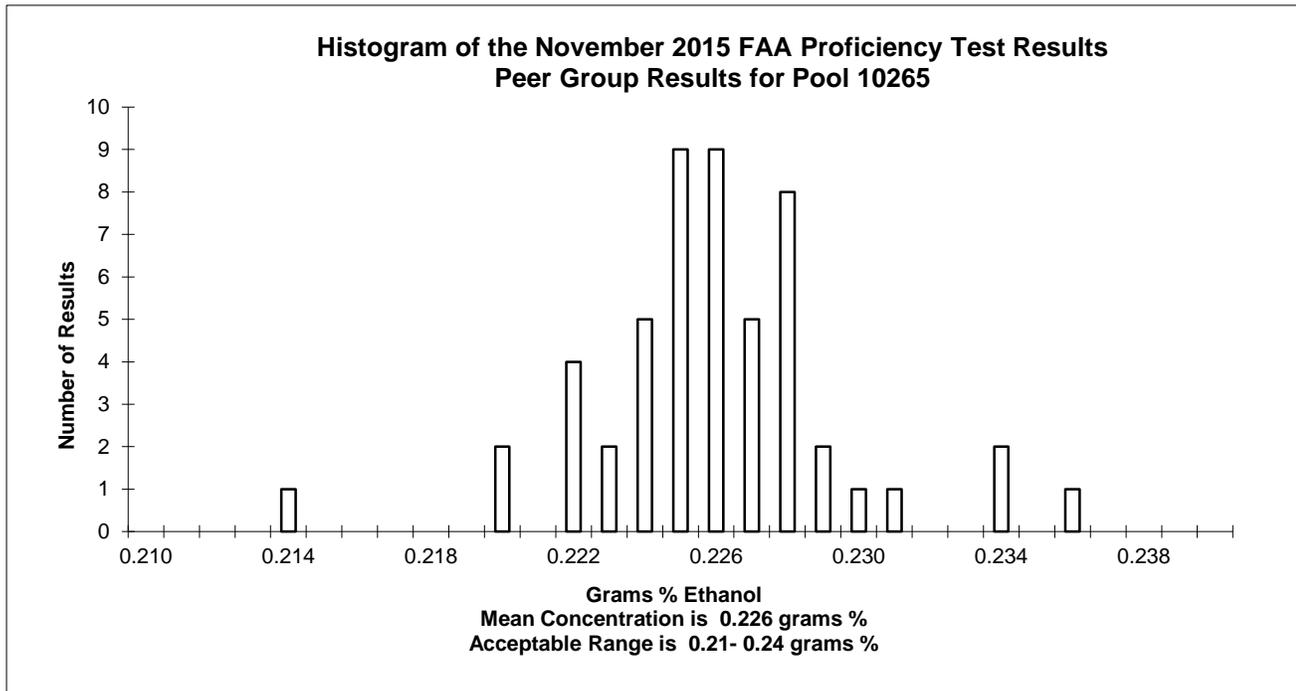
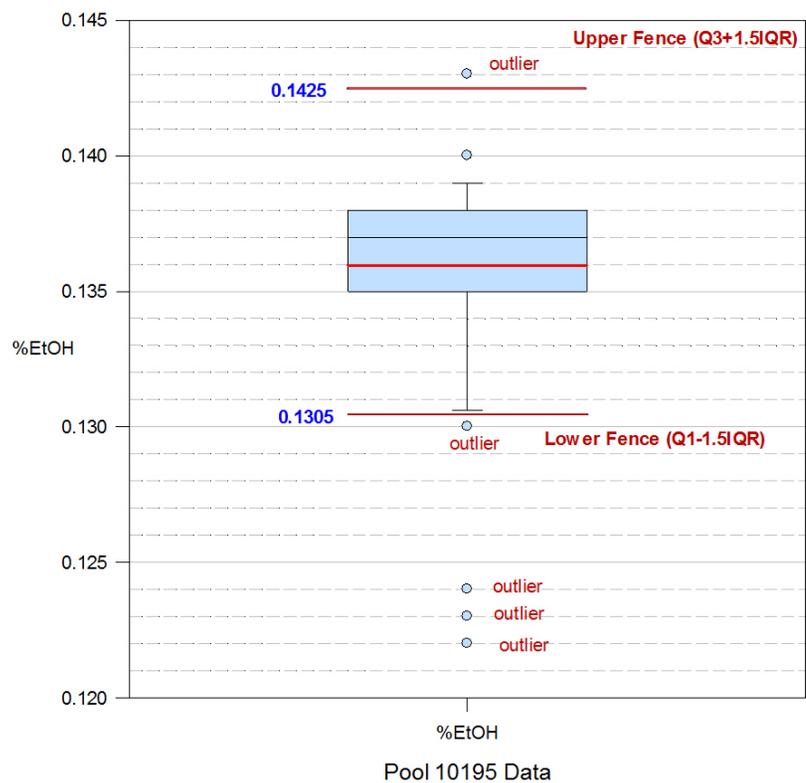


Figure 3 SigmaPlot analysis of pools 10195 & 10265

10195 Box Plot



10265 Box Plot

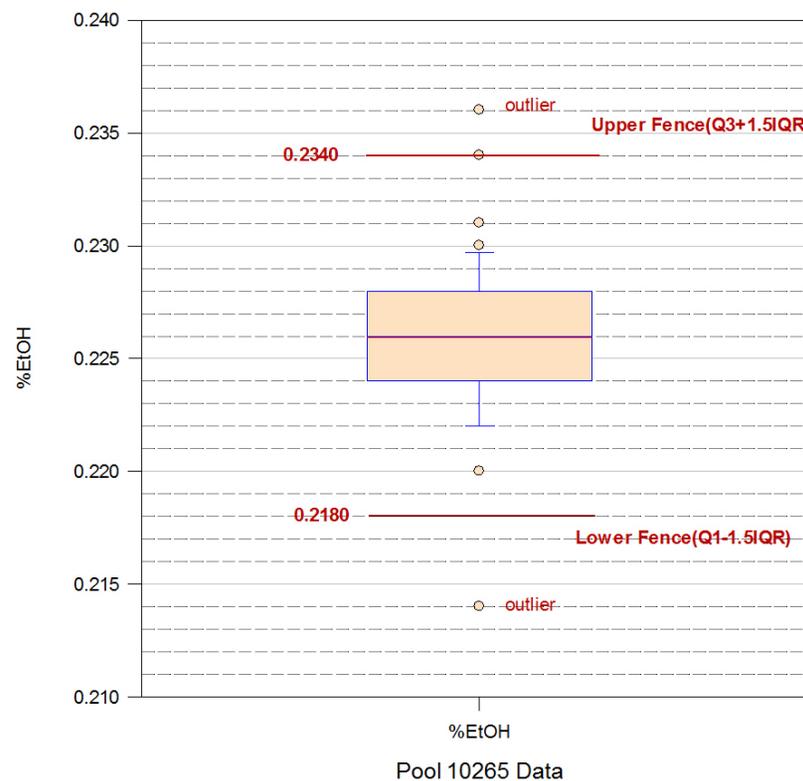


Figure 4

