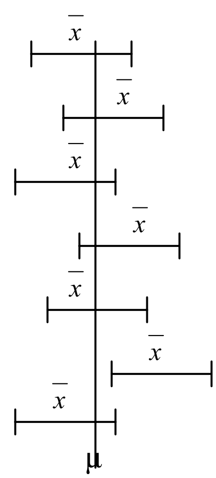
Computing Confidence Intervals in Breath Alcohol Analysis

In forensic breath alcohol analysis we perform duplicate measurements and determine a sample mean $(\bar{x})$, which is used subsequently as an unbiased estimate the individual's true mean (μ) breath alcohol concentration. Our sample mean is a random variable and is simply an estimate of the true mean. We are interested in how good this estimate is. A confidence interval is the best method for determining how well our sample mean estimates the true mean because it establishes an interval within which we can state that we have included (bracketed) μ with some level of probability. An important point to consider is that μ is some theoretically fixed value. On the other hand, our sample mean is a random variable and will vary from duplicate sample to duplicate sample. If we performed 100 duplicate measurements and computed the sample mean in each case, we would obtain a distribution of sample means, which would be approximately normally distributed based on the Central Limit Theorem. Since the sample mean is a random variable, so are the end points for the confidence intervals. The figure below illustrates this concept:



The confidence interval limits is seen to bracket μ in all cases except one. This illustrates that the sample mean along with its limits vary whereas the population mean does not. Properly

interpreting the confidence interval for a population mean would imply that if we were to measure a large number of samples, compute their means and confidence intervals, we would bracket the true population mean 99% of the time (if we determined 99%

confidence intervals).

The standard approach to computing a confidence interval for a population mean is:

$$\bar{x} \pm t_{(1-\alpha/2,\ df)} \frac{S}{\sqrt{n}} \qquad \text{Eq. 1}$$

where: $\bar{x}$ = the sample mean
t = a value selected from the t-table corresponding to a specified level of confidence
1-α/2 = designates the confidence interval width (e.g., α=0.01 would correspond to a 99% confidence interval)
df = degrees of freedom where df = n-1
S = the standard deviation associated with the measurement of a particular breath alcohol concentration and estimated from sample data
n = the number of breath alcohol measurements

In cases where S is based on a large number (n>100) of replicate measurements we will replace S with σ (the true population standard deviation) and replace t with Z (the standard normal variate) and use the standard normal distribution table.

For example, consider an individual producing duplicate breath alcohol results of 0.12 and 0.14 g/210L. The following illustrates the steps for computing the 99% confidence interval:

mean = 0.130 g/210L

$$\sigma = 0.0305\ \bar{x} + 0.0026 = 0.0066 \text{ g/210L} \qquad \text{Eq. 2}$$

where the equation for σ is determined from a large number of duplicate breath alcohol results showing σ to vary with concentration. Now computing the 99% confidence interval we obtain:

$$0.130 \pm 2.576 \frac{0.0066}{\sqrt{2}} = 0.130 \pm 0.012 \qquad \text{Eq. 3}$$

The 99% confidence interval would be reported as:

0.118 to 0.142 g/210L

The confidence interval for a population mean (assuming σ is known based on large n) can be expressed more completely as follows:

$$P\left[\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + \text{Zsub } \alpha/2 \frac{\sigma}{\sqrt{n}}\right] = 0.99 \qquad \text{Eq. 4}$$

Equation 4 is an inequality stating that the probability the upper and lower confidence interval limits will bracket the population mean μ is 0.99. It is important to recognize that the limits of the interval are the random variables and not μ.

In forensic breath alcohol analysis we are more interested in the value of the lower limit than the upper since we are not concerned that the population mean actually exceeds our computed upper limit. Therefore, we are more interested in the following inequality:

$$P\left[\bar{x} \quad Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le +\infty\right] = 0.995 \qquad \text{Eq. 5}$$

or simply:

$$P\left[\bar{x} \quad Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu\right] = 0.995 \qquad \text{Eq. 6}$$

Notice that our probability increases by (100-P)/2 since our interest now is only in one tail of the distribution. Again, Z is used in place of t since our estimate of σ is based on a large number of duplicate breath alcohol values ($n > 1,000$ for most relevant concentrations). For large n (large df) the t distribution converges to the standard normal distribution.

Putting the values from our example in Equation 3 into Equation 6 we obtain:

$$P\left[0.130 \quad 2.576 \quad \frac{0.006}{\sqrt{2}} \le \mu\right] = P\left[0.118 \le \mu\right] = 0.995 \qquad \text{Eq. 7}$$

This is interpreted as meaning the probability that our lower confidence interval limit will be less than or equal to μ is 0.995. Again, the lower limit of 0.118 g/210L is the random variable and not μ.

Probability That μ Exceeds A Specified Lower Limit

There may be occasions when we are interested not only in the value of the lower 99.5% confidence limit but also in the probability that μ exceeds some specific lower value. Consider an individual that provides duplicate breath alcohol samples resulting in 0.100 and 0.110 g/210L. The mean would be 0.1050 g/210L and the standard deviation (σ) determined from Equation 2 would be 0.0058 g/210L. We express this new probability statement as follows:

$$P\left[\mu \geq 0.100 \quad\right] = P\left[0.105 \quad Z \frac{0.0058}{\sqrt{2}} \geq 0.100 \quad\right] \qquad \text{Eq. } 8$$

where: $\mu = 0.105 - Z(0.0058)/(\sqrt{2})$

We now solve for Z and find its associated probability:

$$P\left[0.105 \quad Z \frac{0.0058}{\sqrt{2}} \geq 0.100 \quad\right] = P\left[\frac{0.105 \quad 0.100}{0.0058/ \quad \sqrt{2}} \geq Z\right] = P(1.22 \quad \geq Z)$$

Looking up the value of 1.22 in the standard normal table we see that the probability that Z is less than or equal to 1.22 is 0.889. We therefore conclude that the probability that our lower confidence interval limit is equal to or greater than 0.100 g/210L is 88.9%. This is the same as stating that our lower 88.9% confidence interval limit for our population mean (μ) is 0.100 g/210L.

RGG 3/15/99