Parametric measures of variation around the sample mean

For a dataset of N observations sampled at random from a normally distributed population of interest, the *sample mean* is the average value per observation given by the sum of the data divided by N. Always accompany the sample mean with a description of the sample variation (v or SD) or an inference about the population based on the sample variation (SE or CI).

The sample *variance* (v) is a measure of the spread of data around the sample mean, given by the mean square deviation of the data (the average squared deviation from the mean per degree of freedom[1]). In an Analysis of Variance on two or more samples, the variance equals the MS[Error], which is used to calibrate MS[Test] in calculating the statistic F.

The sample *standard deviation* (SD) equals the square root of v. About 68% of normally distributed data lies within 1 SD either side of the mean, and 95% of values lie within 1.96 SD of the mean. Sample means are often described with their SD, or plotted ± 1SD.

The *standard error of the mean* (SE) is the SD of sample means around the population mean, and so a measure of the variability amongst the means of samples taken from the same population. It equals the square root of the contribution per observation to the sample variance: $\sqrt{(v/N)} = \text{SD}/\sqrt{N}$. An analysis of variance is often illustrated by means ± 1SE.

The *confidence interval* (CI) hosts the range of plausible values of the population mean μ for a population that yields the sample mean and variance. A smaller CI indicates more precision in estimating μ. The Central Limit Theorem dictates that the CI applies even to non-normally distributed populations, given a large enough sample from the population.

The plot below shows a sample mean and 95% CI. Suppose the sample comprises $N = 24$ chicks from a population of interest, and the response y is their change in body mass (g) over the first 12 hours since hatching. The sample mean = 2.60 g shows that these chicks gained weight on average. The CI reveals, however, that a population with normally distributed growth around a mean of $\mu = 0$ would yield sample means at least as deviant as the observed one in more than 5% of samples comprising 24 randomly selected chicks. Likewise, a population with $\mu = 5$ g would yield sample means at least as deviant in > 5% of samples. The CI encompasses the range of plausible values of μ, given only the sample data and the assumption of normality. Thus we fail to reject a null hypothesis $H_0 = 0$ or 5 or anywhere within the CI. In this scenario, we cannot be confident that the conditions experienced by the sample will favour chick growth in the population.

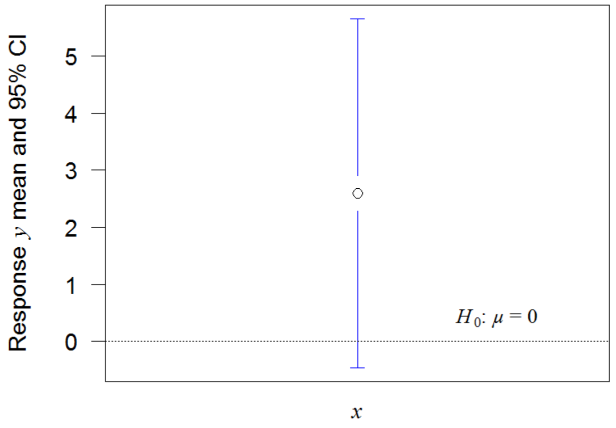

Fig. 1. Sample from a population with unknown μ.

---

[1] For a sample of N data points, the variance has degrees of freedom = $N - 1$. This is because by definition of the sample mean $\bar{y}$, the sum $\sum_i (y_i - \bar{y}) = 0$; thus if we know $N - 1$ of the deviations, we can calculate the last one, meaning that there are only $N - 1$ freely varying deviations.

Formulae for parametric measures of variation

Sample mean: $\bar{y} = \sum_{i=1}^{N} y_i \Big/ N$

Sample variance: $v = \sum_{i=1}^{N} \left(y_i - \bar{y} \right)^2 \Big/ \left(N - 1 \right)$

Standard deviation: $\mathrm{SD} = \sqrt{v}$

Standard error of the mean: $\mathrm{SE} = \mathrm{SD} \big/ \sqrt{N}$

Confidence interval: CI = $\bar{y} \pm t_{[\alpha]\,N-1} \cdot \mathrm{SE}$

One-sample test of $H_0 = \mu$: Student's $t = (\bar{y} - \mu)/\mathrm{SE}$

The confidence interval, CI, for a threshold Type-I error α, lies either side of the sample mean between limits $\bar{y} \pm t_{[\alpha]\,N-1} \cdot \mathrm{SE}$, where the quantile $t_{[\alpha]\,N-1}$ is the critical value of the Student's t distribution at two-tailed α for $N - 1$ degrees of freedom [given by the R command: `qt(1-alpha/2, N-1)`]. For a very large sample, the 95% CI are well approximated by 1.96×SE. For the Fig.-1 data, $\bar{y} = 2.60$ g, SE = 1.477, $N = 24$, and $t_{[0.05]\,23} = 2.069$; so the 95% CI = 2.60 ± 3.05 g.

The one-sample Student's t statistic tests the compatibility of the data with a refutable null hypothesis: $H_0 = \mu$. The data give $t = (\bar{y} - \mu)/\mathrm{SE}$, with $N - 1$ degrees of freedom. The probability, P, of Type-I error (rejecting a true null hypothesis) can be found by evaluating t against its α-quantile, $t_{[\alpha]\,N-1}$ for a two-tailed test (obtained either with the R command given above, or from a table of α-quantiles of the Student's t distribution). If t exceeds the critical value, then $P < \alpha$ and we reject H_0. Alternatively, P is obtained directly with the R command: `2*(1 - pt(abs(t), N-1))`. For example, an analysis of the Fig.-1 data fails to refute the null hypothesis of zero growth ($t_{23} = 1.76$, $P = 0.09$). The value of P is the probability of data at least as deviant given the null hypothesis, and thus the probability of making a mistake by rejecting a true H_0. In other words, a normally distributed population with mean $\mu = 0$ g has 9% probability of yielding a sample mean at least as deviant as the observed $\bar{y} = 2.60$ g. At > 5%, this is too high a probability for us to reject the possibility of zero growth in the population. The result accords perfectly with the inference we drew from the 95% CI (page 1 above and Fig. 1).

R commands for plotting Fig. 1

```
# Plot of a single sample mean with 95% CI.
rm(list = ls())
##################################################################
library(gplots) # *** requires installation of package 'gplots' ***
mu <- 0         # the null hypothesis for the value of the population mean
alpha <- 0.05   # critical threshold for Type-I error
y <- c(-3.4076, -8.5488, -1.8395, -5.6626, -2.8127, -4.6421,
       -1.6908, -4.7169, -0.6191, 1.2085, 5.1147, 7.2654, 4.4188,
       6.3951, 0.5986, -4.5055, 13.322, 17.0426, 14.66, 16.1283,
       8.5927, 2.2129, 1.8934, 2.0203) # the data
##################################################################
N <- length(y) ; DF <- N - 1
x <- c(rep("1", N))        # set x = 1 for all values of this single sample
SE <- sqrt(var(y)/N)       # standard error of the mean: SE = sqrt(v/N)
CL <- SE*qt(1-alpha/2, DF) # confidence limits either side of mean
setcex <- 1.8              # set the font size for labels
par(cex = setcex, mar = c(4, 4, 1, 2) + 0.1) # define plot margins
boxplot(y) # Box and whisker plot of the data
# Plot the mean and CI using 'plotmeans' from the 'gplots' library
plotmeans(y ~ x, p = 1-alpha, las = 1, xlab = "", ylab = "", n.label = FALSE, xaxt = "n")
# Add  reference line, axis labels, and legend
windowsFonts(A = windowsFont("Times"),
             B = windowsFont("Arial"),
             C = windowsFont("Cambria"))
abline(h = mu, lty = 3) # plot a reference line for the null hypothesis
if (mu > (mean(y)+qt(1-alpha/2,DF))*0.9) {ps = 1} else {ps = 3} # text below/above line
text(1.3,mu,substitute(paste(italic("H"),""[0],":"~italic("\u03bc")~"= ",v), list(v=mu)),
     pos=ps, cex=setcex/2, family = "A")
mtext("x", font = 3, side = 1, line = 0.5,
      las = 1, cex = setcex, family = "A")
mtext(substitute(paste("Response    mean and ",v,"% CI"), list(v=100*(1-alpha))),
      side = 2, line = 2.5, las = 0, cex = setcex, family = "B")
mtext(expression("    "~italic("y")~"                  "),
      side = 2, line = 2.5, las = 0, cex = setcex, family = "A")
mtext(expression(~
      bold("Fig. 1.")~" Sample from a population with unknown"~italic("\u03bc .")),
      outer = TRUE, side = 1, line = -1.5, cex = setcex, family = "C")
# Report statistics
writeLines(sprintf("Sample mean = %.2f g.",mean(y))) # mean (with trailing zero)

## Sample mean = 2.60 g.

t <- (mean(y)-mu)/SE ; P <- 2*(1-pt(abs(t),DF)) # Student's t test
tcrit <- qt(1-alpha/2, DF) # critical t at alpha for a two-tailed test
if (abs(t) > tcrit) {result <- "Reject"} else {result <- "Fail to reject"}
writeLines(c(result, " H0: mu = ", mu, sprintf(" (t = %.3f, ",t),
             sprintf("DF = %.0f, ",DF), sprintf("P = %.3f; ",P),
             sprintf("t[alpha=%.2f] ",alpha), sprintf("= %.3f).",tcrit)
             ),sep = "")

## Fail to reject H0: mu = 0 (t = 1.762, DF = 23, P = 0.091; t[alpha=0.05] = 2.069).

#
par(par(no.readonly = TRUE))
```