Statistics for Computing MA4413

Lecture 11

Sum / Difference of Independent Normal Variables and Calculating Normal Limits

Kevin Burke

kevin.burke@ul.ie

Sum / Difference of Normal Variables

If $X_1 \sim \text{Normal}(\mu_1, \sigma_1)$ and $X_2 \sim \text{Normal}(\mu_2, \sigma_2)$ are **independent normal variables** then the **sum** is

$$\boxed{X_1 + X_2 \sim \text{Normal}(\mu = \mu_1 + \mu_2,\ \sigma = \sqrt{\sigma_1^2 + \sigma_2^2}\,)},$$

and the **difference** is

$$\boxed{X_1 - X_2 \sim \text{Normal}(\mu = \mu_1 - \mu_2,\ \sigma = \sqrt{\sigma_1^2 + \sigma_2^2}\,)}.$$

Note that in *both* cases $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$.

Example: Batteries

Let the voltages for two batteries be $X_1 \sim \text{Normal}(\mu_1 = 6, \sigma_1 = 0.1)$ and $X_2 \sim \text{Normal}(\mu_2 = 6, \sigma_2 = 0.1)$ where X_1 and X_2 are independent.

Let's assume that a 12V battery is made up of two of these batteries. Let Y represents the total voltage:

$$
\begin{aligned}
Y = X_1 + X_2 &\sim \text{Normal}(\mu = 6 + 6,\ \sigma = \sqrt{0.1^2 + 0.1^2}\,) \\
&\sim \text{Normal}(\mu = 12,\ \sigma = 0.1414).
\end{aligned}
$$

We can then calculate probabilities as before, e.g.,

$$
\Pr(Y > 12.15) = \Pr(Z > \tfrac{12.15-12}{0.1414}) = \Pr(Z > 1.06) = 0.1446.
$$

Question 1

Let X_1 represent the time it takes a person to complete a particular task where $X_1 \sim \text{Normal}(\mu = 45, \sigma = 1)$. Another individual's time is $X_2 \sim \text{Normal}(\mu = 44, \sigma = 1.5)$. Let $Y = X_1 - X_2$.

a) What is the distribution of Y?

b) What is the probability that person 1 finishes first?

c) What is the probability that person 2 finishes first?

d) What is the probability that the winner finishes at least 2 seconds before the other person?

95% Limits

For a variable $X \sim \text{Normal}(\mu, \sigma)$, it is often of interest to calculate **limits** x_1 and x_2 such that

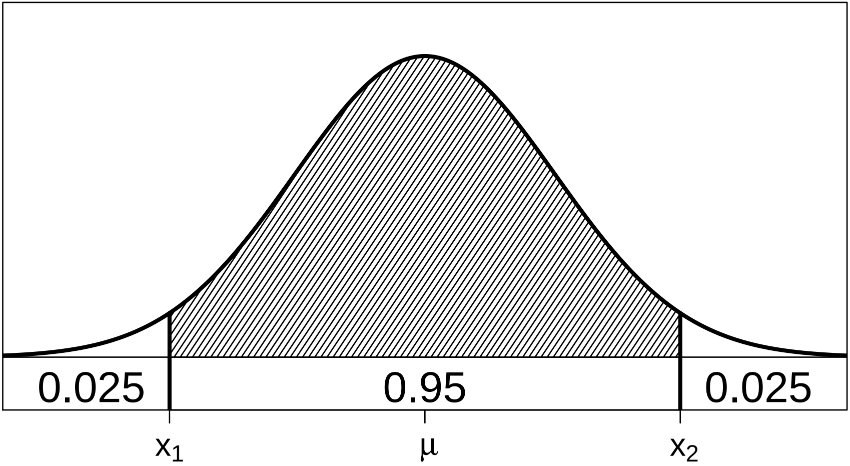
$$\Pr(x_1 < X < x_2) = 0.95.$$

More specifically, this interval is constructed to cover the **central 95%** of the distribution.

Thus 5% of the distribution remains:

- 2.5% in the lower tail (below x_1)
- 2.5% in the upper tail (above x_2)

95% Limits



Example: Salary

We return to the salary example from the previous lecture where $X \sim \text{Normal}(\mu = 30, \sigma = 4)$.

We will now calculate the 95% salary limits:

lower tail	**upper tail**
$\Pr(X < x_1) = 0.025$	$\Pr(X > x_2) = 0.025$
$\Pr(Z < \frac{x_1 - 30}{4}) = 0.025$	$\Pr(Z > \frac{x_2 - 30}{4}) = 0.025$
$\Pr(Z > -\frac{x_1 - 30}{4}) = 0.025$	

We find that the z score which corresponds to $\Pr(Z > z) = 0.025$ is $z = 1.96$ (from the tables).

Example: Salary

lower limit

$$-\frac{x_1-30}{4} = 1.96$$

$$\frac{x_1-30}{4} = -1.96$$

$$x_1 - 30 = -1.96(4)$$

$$x_1 = 30 - 1.96(4)$$

$$x_1 = 22.16$$

upper limit

$$\frac{x_2-30}{4} = 1.96$$

$$x_2 - 30 = 1.96(4)$$

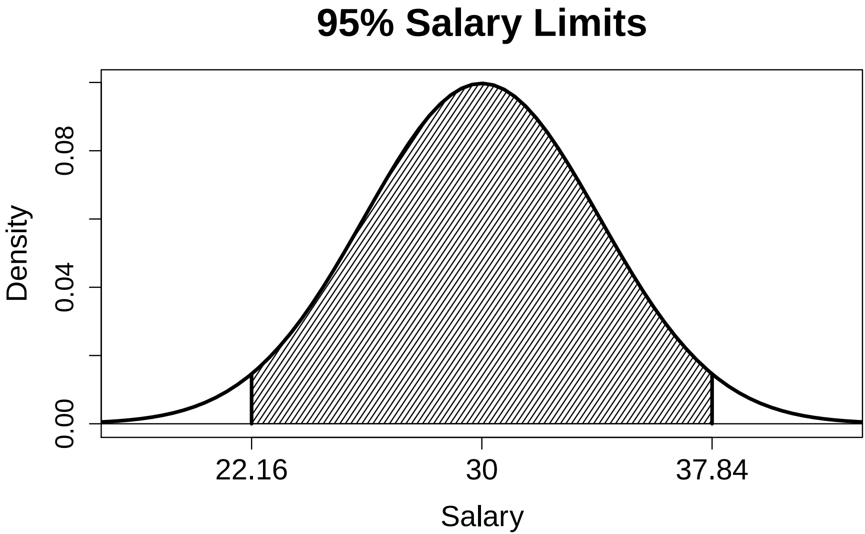
$$x_2 = 30 + 1.96(4)$$

$$x_2 = 37.84$$

In words, the central 95% of salaries lie in the interval $[22.16,\ 37.84]$.

Note: this interval can be written $30 \pm (1.96 \times 4)$. *(more on this later)*

Example: Salary



$(1-\alpha)100\%$ Limits

We can calculate the limits for other percentage values:

$$\Pr(x_1 < X < x_2) = 1 - \alpha,$$

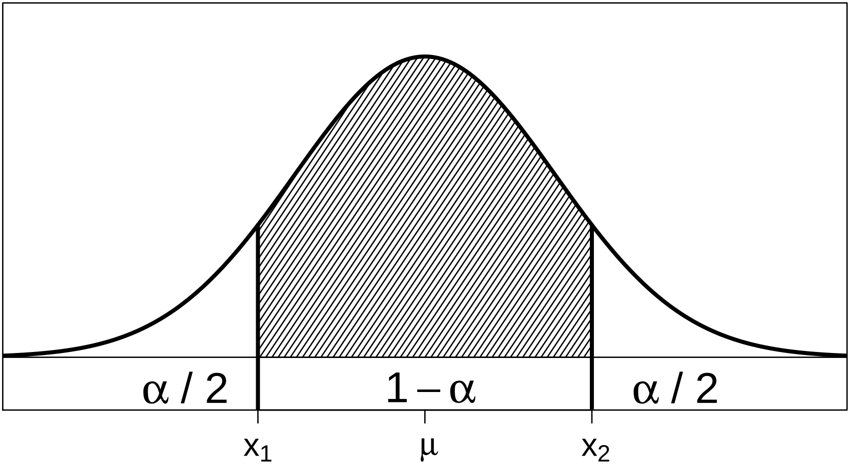
where x_1 and x_2 are chosen so that the central probability of $1-\alpha$ is covered (α is the Greek letter "alpha").

Thus the probability α remains:

- $\alpha/2$ in the lower tail (below x_1)
- $\alpha/2$ in the upper tail (above x_2)

Example: $\alpha = 0.01 \Rightarrow (1 - 0.01) \times 100\% = 99\%$ limits with $\alpha/2 = 0.005$ remaining in each tail (i.e., 0.5%).

$(1-\alpha)100\%$ Limits



Question 2

We continue with salary $X \sim$ Normal($\mu = 30, \sigma = 4$).

a) Calculate the limits within which the central 99% of salaries lie.

Example: Salary

We noted that the 95% salary limits can be written in the form:

$$30 \pm (1.96 \times 4)$$

We now let $z_{0.025} = 1.96$ since it is the z score corresponding to a probability of 0.025 in the tables, i.e., $\Pr(Z > 1.96) = 0.025$.

Thus, the 95% limits are:

$$30 \pm (z_{0.025} \times 4)$$

Similarly, the 99% limits are:

$$30 \pm (z_{0.005} \times 4)$$

$\Rightarrow$ The $(1 - \alpha)100\%$ limits are:

$$30 \pm (z_{\alpha/2} \times 4)$$

Constructing Limits

For a general Normal(μ, σ) distribution, the interval

$$\boxed{\mu \pm z_{\alpha/2}\ \sigma}$$

contains $(1 - \alpha)100\%$ of the distribution. This fact can be stated mathematically as $\Pr(\mu - z_{\alpha/2}\ \sigma < X < \mu + z_{\alpha/2}\ \sigma) = 1 - \alpha$.

In practice, we do not need to go through the probability arguments of the previous slides every time.

Simply look up the $z_{\alpha/2}$ score in the tables and use the above formula.

Commonly Used Limits

It is very common to compute the following:

%	α	$z_{\alpha/2}$	Interval
90%	0.10	$z_{0.05} = 1.64$	$\mu \pm 1.64\,\sigma$
95%	0.05	$z_{0.025} = 1.96$	$\mu \pm 1.96\,\sigma$
99%	0.01	$z_{0.005} = 2.58$	$\mu \pm 2.58\,\sigma$

Of course, any interval can be calculated, e.g., $\mu \pm 1.28\,\sigma$ gives the 80% limits, $\mu \pm 0.67\sigma$ gives the 50% limits etc.

R Code

We can look up $z_{\alpha/2}$ scores in the normal tables. We can also do this using `qnorm`:

Examples:

`qnorm(0.05,mean=0,sd=1,lower=F)` gives `1.644854`.

`qnorm(0.025,mean=0,sd=1,lower=F)` gives `1.959964`.

`qnorm(0.005,mean=0,sd=1,lower=F)` gives `2.575829`.

Compare the above to the previous slide.

R Code

Note that we can also construct the intervals directly using `qnorm`:

Examples:

`qnorm(0.025,mean=30,sd=4,lower=T)` gives `22.16014`.

`qnorm(0.025,mean=30,sd=4,lower=F)` gives `37.83986`.

Compare this with slide 8.