

DAILY ASSESSMENT FORMAT

Date:	30/07/2020	Name:	Prajwal Kamagethi Chakravarti P L
Course:	Coursera	USN:	4AL17EC073
Topic:	• Basic Statics	Semester & Section:	6 & B
Github Repository:	https://github.com/alvas-education-foundation/Prajwal-Kamagethi.git		

SESSION DETAILS

Session images

The screenshot shows a web browser window displaying a DataCamp exercise titled "Practise With Sets III". The exercise is part of a course on "Basic Statistics". The instructions state: "Now let's pretend that you have 6 alcohol drinkers and 4 non-alcohol drinkers in a room. Without knowing who is who, you pick three in sequence to take part in your study. Let's consider the probability of various outcomes." The exercise is worth 100 XP. The instructions list: "Answer the questions in your script by writing out your calculations in full below" and "Do not simplify your fractions". There is a "Take Hint (-30 XP)" button. The main area shows an R script editor with the following code:

```
1 # Calculate the probability that the first two people drink alcohol
2
3 # Calculate the probability that at the first and third person drink
  alcohol
4
5 # Calculate the complement of the the probability that none of the
  people drink
6
```

Below the script editor is an "R Console" area. At the bottom right of the script editor are buttons for "Run Code" and "Submit Answer". The browser's address bar shows the URL: campus.datacamp.com/courses/basic-statistics/lab-3-probability?ex=14. The Windows taskbar at the bottom shows the time as 09:56 on 29-07-2020.

Report:

Low statistical power:

Power is the probability of correctly rejecting the null hypothesis when it is false (inverse of the type II error rate). Experiments with low power have a higher probability of incorrectly accepting the null hypothesis—that is, committing a type II error and concluding that there is no effect when there actually is (I.e. there is real covariation between the cause and effect). Low power occurs when the sample size of the study is too small given other factors (small effect sizes, large group variability, unreliable measures, etc.).

Violated assumptions of the test statistics:

Most statistical tests (particularly inferential statistics) involve assumptions about the data that make the analysis suitable for testing a hypothesis. Violating the assumptions of statistical tests can lead to incorrect inferences about the cause-effect relationship. The robustness of a test indicates how sensitive it is to violations. Violations of assumptions may make tests more or less likely to make type I or II errors.

Dredging and the error rate problem:

Each hypothesis test involves a set risk of a type I error (the alpha rate). If a researcher searches or "dredges" through their data, testing many different hypotheses to find a significant effect, they are inflating their type I error rate. The more the researcher repeatedly tests the data, the higher the chance of observing a type I error and making an incorrect inference about the existence of a relationship.

Unreliability of measures:

If the dependent and/or independent variable(s) are not measured reliably (i.e. with large amounts of measurement error), incorrect conclusions can be drawn.

Restriction of range:

Restriction of range, such as floor and ceiling effects or selection effects, reduce the power of the experiment, and increase the chance of a type II error.^[5] This is because correlations are attenuated (weakened) by reduced variability (see, for example, the equation for the Pearson product-moment correlation coefficient which uses score variance in its estimation).

