Statistical Methods II

Quiz 01

Name: Date:

Welcome to the first quiz. It is a practice quiz designed for you to see what types of information I expect you to have for the quiz. You have five (5) minutes to answer these 10 questions. The total number of points is 10.

⋆ ⋆ ⋆

Distributions
Let us suppose $X \sim \mathcal{N}(\mu = 1, \sigma^2 = 3)$, $Y \sim \mathcal{N}(\mu = 3, \sigma^2 = 2)$, and $Z \sim \mathcal{N}(0, 1)$. Answer the following questions.

1 (1 point)
What is the mean of $X + Y$?

> **Solution:**
>
> Means always add. Therefore, the mean of $X + Y$ is $1 + 3 = 4$.

2 (1 point)
What is the variance of $X + Z$?

> **Solution:**
>
> Assuming X and Z are independent (as the problem implies), variances also add. Therefore, $\mathbb{V}[X + Z] = 3 + 1 = 4$.

3 (1 point)
What is the distribution (including mean and variance) of X+Y?

> **Solution:**
>
> First, we know that the sum of Normal distributions is Normal (one of the nice things about it). Second, we know that the means add as do the variances. Thus,
>
> $$X + Y \sim \mathcal{N}(\mu = 4, \sigma^2 = 5)$$

4 (1 point)
What is the distribution (including mean and variance) of $3X + 2Y$?

Solution:

Again, we know that the sum of two Normal distributions is also a Normal distribution. We also know that the expectation operator is linear, thus, the expected value of $3X + 2Y$ is

$$\mathbb{E}[3X + 2Y] = 3\mathbb{E}[X] + 2\mathbb{E}[Y] = 3 \times 1 + 2 \times 3 = 9$$

Unfortunately, variances are not linear operators. (Covariance, however, is a bilinear operator.) We do recall, however, that $\mathbb{V}[aX] = a^2\mathbb{V}[X]$. Thus, we have

$$\mathbb{V}[3X + 2Y] = 9\mathbb{V}[X] + 4\mathbb{V}[Y] = 35$$

Thus, we conclude that

$$3X + 2Y \sim \mathcal{N}(\mu = 9, \sigma^2 = 35)$$

True and False
Please write 'TRUE' in the space provided if the statement is true; 'FALSE', if the statement is not true.

5 (1 point)
Two of the assumptions of the t-test are that the measurements are Normally distributed and that the measurements are independent.

5. _____

Solution:

TRUE. When it was first presented to you, we also had to assume the variances were equal. However, there are t-tests that do not require equal variances.

The assumptions come from the structure of the t-statistic:

$$t := \frac{\overline{X} - \overline{Y}}{\sqrt{\sigma^2/n}}$$

Only if the numerator is distributed Normal will this statistic be distributed t.

6 (1 point)
The null hypothesis of a two-sample t-test is that the means of the two populations are the same.

6. _____

Solution:

TRUE. From the definition of the t-statistic, if the null hypothesis is correct, the statistic will have a value close to zero.

7 (1 point)
If the calculated p-value is less than your chosen alpha, you should reject the null hypothesis.

7. _____

Solution:

TRUE. The p-value is the probability of getting results this extreme (or more so) if the null hypothesis is correct. Thus, small p-values suggest that the null hypothesis is not supported by the data. R. A. Fisher chose $\alpha = 0.05$ to represent the rate of being wrong he was willing to be (5% of the time when rejecting the null hypothesis).

8 (1 point)
Power is a measurement of how well the test is able to distinguish between two competing hypotheses.

8. _____

Solution:

TRUE. Recall that a powerful test will be able to determine from which distribution a sample comes.

R Questions

What do the following commands do? Make sure you make it clear that you know what the function and its given parameters do.

9 (1 point)

```
read.csv("dogfight.csv", header=TRUE)
```

Solution:

This statement will read in the dataset with filename "dogfight.csv" and treat the first row as the variable names.

10 (1 point)

```
t.test(var1,var2, var.equal=TRUE)
```

Solution:

This statement will perform an equal-variance t-test, statistically comparing the means of `var1` and `var2`.