

The following questions are to help you verify that you studied **all** of the material in the **Making Inference** portion of the **Statistics Notebook**. Remember, real learning only happens as you take time to study the material.

True or False. Inference is the process of deciding whether patterns and trends in a sample of data from a population can be assumed to be true for the full population.

- ☒ True
- ☐ False

True or False. A pattern can appear one way in a sample of data, but be completely different for the full population. That is why we use statistical inference to decide when it is safe (and when it is not) to conclude that a pattern in a sample holds in the full population.

- ☒ True
- ☐ False

True or False. If a researcher was using a significance level of 0.05 and obtained a p-value of 0.419, they can safely conclude that the null hypothesis is false and that the alternative hypothesis is the truth.

- ☐ True
- ☒ False

True or False. In a typical U.S. Court of Law, the null hypothesis is that the person on trial is innocent.

- ☒ True
- ☐ False

Thus, convicting an innocent man of a crime would be an example of

- ☐ A correct decision.
- ☒ A Type I Error.
- ☐ A Type II Error.

And letting a guilty man go free would be an example of

- ☐ A correct decision.
- ☐ A Type I Error.
- ☒ A Type II Error.

The p-value is the most commonly used method of deciding when to reject a null hypothesis. The things needed to compute a p-value are:

- ☒ A test statistic.
- ☒ A probability distribution of the test statistics that would be possible if the null hypothesis were true.

The parametric distributions this course will teach you about are:

- ☒ The Normal Distribution
- ☐ The Weibull Distribution
- ☒ The t Distribution
- ☒ The F Distribution
- ☒ The Chi-squared Distribution
- ☐ The Uniform Distribution

In which distribution would an observation of $x=5$ be the most likely to occur?

Copy and paste the following R Codes in your Console for a graph that goes along with this question:

```
curve(dnorm(x,-5,2), -12, 20, col="skyblue", lwd=2, ylab="", n=1000, ylim=c(0,.81))
curve(dchisq(x,5), from=0, add=TRUE, col="firebrick", lwd=2, n=1000)
curve(dt(x,5), add=TRUE, col="skyblue4", lwd=2, n=1000)
curve(df(x,10,10), from=0, add=TRUE, col="salmon", lwd=2, n=1000)
legend("topright", legend=c("Normal(-5,2)", "Chi-squared(8)", "t(5)", "f(2,5)"), lwd=2, lty=1,
col=c("skyblue", "firebrick", "skyblue4", "salmon"), bty='n')
```

- ☐ A normal distribution with $\mu = -5$ and $\sigma = 2$.
- ☒ A chi-squared distribution with $p = 8$.
- ☐ A t distribution with $p = 5$
- ☐ An F distribution with $p_1 = 2$ and $p_2 = 5$.

In your R Console, copy and paste the commands:

```
> curve(dnorm(x, -5, 2), from=-12, to=2, lwd=2, col="skyblue", ylab="Normal Distribution with Mean of -5 and Standard Deviation of 2")
> abline(h=0, v=-5, lty=2)
```

Which of the following set of sample values are most likely to be obtained if a sample of $n = 5$ observations is sampled from the normal distribution drawn by the above code in R?

- ☐ -12, -8, -10, -11, -7
- ☐ 3, 8, 1, 15, 2
- ☒ -8, 0, -5, -6, -3