1. For each of the following statements, indicate whether the statement in **bold** is true or false by circling the appropriate choice.

(b), (d) are required

(a) Suppose we wish to use the bootstrap to make a confidence interval for the mean of a population with an unknown distribution. We take a sample from the population of size n. We use the statistic $\hat{t} = \frac{\bar{x}^* - \bar{x}}{\frac{s^*}{\sqrt{n}}}$, where $\bar{x}$ is the sample mean of the original sample, $\bar{x}^*$ is the sample mean of the resampled data, and s^* is the sample standard deviation of the resampled data. Suppose we resample B times. **The larger the value of B, the more like a normal the distribution of $\hat{t}$ will be, because of the central limit theorem.** TRUE FALSE

(b) **The p-value is the probability that the null hypothesis is true.** TRUE FALSE

(c) **When performing a linear regression, the y variable (response variable) must be normally distributed in order for the t-test for the slope to be valid.** TRUE FALSE

(d) **A p-value smaller than α calls for rejecting the null hypothesis.** TRUE FALSE

ANSWER:

(a) FALSE

(b) FALSE

(c) FALSE

(d) TRUE

2. A certain chemical reaction was run three times at each of three temperatures. The yields, expressed as a percent of a theoretical maximum, were as follows:

Temperature (°C)	Yields			Row Mean
70	81.2	82.6	77.4	80.4
80	93.3	88.9	86.0	89.4
90	87.8	89.2	88.5	88.5
				Overall Mean = 86.1

Consider testing $H_0 : \mu_{70^\circ} = \mu_{80^\circ} = \mu_{90^\circ}$. Suppose the assumptions of ANOVA are met.

(a) Find the treatment sum of squares.
ANSWER: $3 * [(80.4 - 86.1)^2 + (89.4 - 86.1)^2 + (88.5 - 86.1)^2] = 147.42$

(b) The error sum of squares is about 42.48. Find the error mean square.
ANSWER: MSE $= \frac{42.48}{9-3} = 7.08$

(c) The F statistic is about 10.4. Find the p-value (providing a small range is ok) and make a conclusion in the context of the problem.
ANSWER: p-value $= P(F_{3-1,9-3} > 10.4) = P(F_{2,6} > 10.4) =$ between .01 and .05
The data are strong evidence that the population mean yields are not the same across the three temperatures.

4. A roller coaster holds 60 riders. It's maximum safe total rider weight is 12,000 pounds. The weights of adult U.S. men have mean 194 and standard deviation 68 pounds. If a random sample of 60 men ride the coaster, what is the probability the maximum safe weight will be exceeded?

ANSWER:

Let X_i = weight of i^{th} randomly-chosen man; X_i has $\mu = 194$ and $\sigma = 68$. CLT says $\bar{X} \sim N(194, 68^2/60) = N(194, 8.779^2)(\approx)$.

$P(\sum_{i=1}^{60} X_i > 12000) = P(\frac{1}{60}\sum X_i > \frac{12000}{60}) = P(\bar{X} > 200) = P(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} > \frac{200-194}{8.779}) = P(Z > .68) = P(Z < .68) = .2483$

5. The average numbers of defectives produced by two press machines are given below:

Population	Observation											
Machine 1	10.5	7.1	5.4	11.3	4.2	11.5	3.6	5.4	5.4	5.2	10.3	9.5
Machine 2	5.4	9.9	5.9	5.3	6.2	6.3	9.6	10	9.4	7.9	6.8	8.5

We are interested in whether Machine 1 is better than Machine 2 in the sense of the number of defectives.

(a) Assume the two populations are normal and the two population variances are equal. Let μ_1 be the population mean for Machine 1 and μ_2 be the population mean for Machine 2. Perform the test at $\alpha = 0.01$.

(b) Find the p-value.

(c) Find the 99% confidence interval for $\mu_1 - \mu_2$.

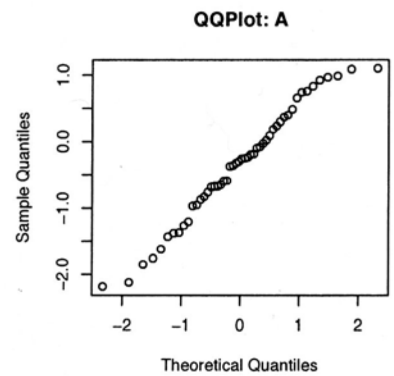
ANSWER:

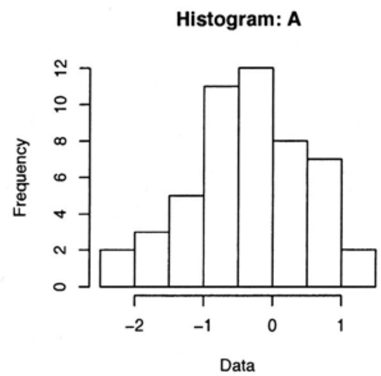
(a) $H_0 : \mu_1 - \mu_2 = 0$; $H_a : \mu_1 - \mu_2 < 0$; $t_{obs} = -0.1497$, rejection region is $t < -t_{22,0.01} = -2.51$. We fail to reject the null and there is not sufficient eveidence that Machine 1 is better than Machine

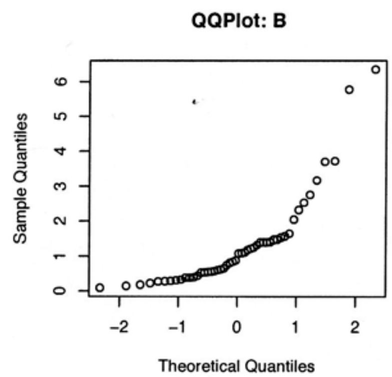
(b) p-value$=P(t < -0.1497) = 0.4412$.

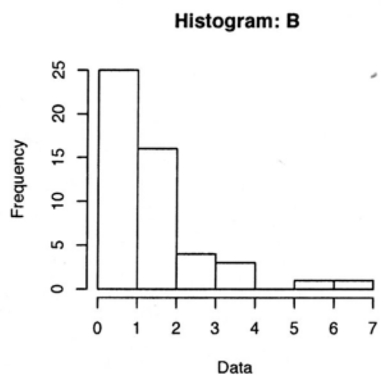
(c) (-2.974, 2.674).

7. Consider these graphical representations of samples from two different populations:









Answer the following questions:

a) Which histogram (A or B) pertains to QQPlot A? Why?
ANSWER: A because the linear QQplot suggests a normal distribution, and Histogram A looks closer to normal than Histogram B.

b) Describe Histogram B in terms of skewness.
ANSWER: Histogram B is right-skewed.

8. Two independent random variables X and Y have normal distributions. $X \sim N(10, 3^2)$ and $Y \sim N(20, 4^2)$. Calculate, with the standard normal table,

$$P(X - Y < 0)$$

Hint: Find the distribution of $X - Y$ and then standardize. (Second hint, added 8/1/16: If X and Y are both normally distributed, then so is the linear combination $X - Y$.)

ANSWER: Let $A = X - Y$. By the formula for linear combination of normal RVs (8/1/16: I didn't include this formula in summer lectures, so I added the "Second hint," above), we have

$$A \sim N(-10, 25).$$

Then

$$\begin{aligned} P(X - Y < 0) &= P(A < 0) \\ &= P(Z < \frac{0 - (-10)}{5} = 2) \\ &= .977 \end{aligned}$$

9. Consider two distinct coins which were found backstage after a magic show. You wish to test whether these two coins have the same probability of landing on Heads with $\alpha = .02$. To this end, you flip the first coin 12 times, giving 3 heads, and the second coin 15 times, giving 4 heads. Ignore verification of assumptions for this question. Answer the following questions:

   a) Write down null and alternative hypotheses which reflect this situation in terms of π_1 and π_2, the true probability of heads for the first and second coin, respectively.
   ANSWER: $\mathcal{H}_0 : \pi_1 - \pi_2 = 0$ and $\mathcal{H}_A : \pi_1 - \pi_2 \neq 0$.

   b) Build a rejection region and decide the test. Interpret the decision in the context of the problem.
   ANSWER: Use $\hat{\pi} = (3 + 4)/(12 + 15) = .259$. Then

$$z = \frac{3/12 - 4/15 - 0}{\sqrt{.259 \times (1 - .259) \times (1/12 + 1/15)}} = -0.09823$$

   The rejection region is $|z| > z_{.01} = 2.32$. Our z is not in this region, so we do not reject $\mathcal{H}_0$. We do not have significant evidence that the coins have different probabilities of Heads.