Contest Quiz 3
Question Sheet

In this quiz we will continue to work with the bank data set, available at:

   http://thiloklein.de/R/bank.

NOTE: Please round your results to two decimal places. Do not round any interim results. EXAMPLE: If your unrounded solution is 0.13897439, drop all decimal places except the first three. This leaves you with 0.138. If the third decimal place is 5 or above (as is the case here), round up. This gives 0.14.

Question 1

I) Calculate the probability that a Normally distributed random variable with mean of 7 and standard deviation of 3 takes the value $x = 5$.

II) Calculate the probability that a Gamma distributed random variable with shape of 10 and scale of 1 has value less than 10.

III) Calculate the value of x, such that the probability that an exponentially distributed random variable with rate of 2 has value greater than x is equal to 0.75.

Question 2

Suppose you are told that nails produced by a factory have an average length of 35mm with a standard deviation of 2mm. Suppose you pick 20 nails at random, and find that the average length in your sample is 33mm.

I) If we wish to perform a hypothesis test to determine whether there is evidence to suggest the average length deviates from 35mm, which test would be most appropriate?
     (a) Z-test (b) t-test

II) Calculate and report the z-score or t-statistic (depending on your answer to I), and report its value to 2 decimal places.

III) Find the 95% critical region for a one sided hypothesis test with H0: mean=35mm; H1: mean<35mm. Which of these is correct?
     (a) (35.74, Inf) (b) (-Inf, 34.26) (c) (-Inf, 34.12) and (35.87, Inf) (d) None of the above

IV) What do you conclude?

     (a) reject the null, accept the alternative hypothesis at the 95% confidence level
     (b) accept the null, reject the alternative hypothesis at the 95% confidence level

Question 3

It has been suggested that the average salary in the bank is $35,000.

I) Calculate and report the appropriate test statistic (z-score or t-statistic) to check whether this is likely to be the case at the 95% confidence level.

II) Construct confidence intervals for the population mean. Which of the following is correct?
(a) (31253, 36386) (b) (32232, 37288) (c) (30127, 35492) (d) (32425, 35545)

III) What do you conclude?

    (a) reject the null, accept the alternative hypothesis at the 95% confidence level

    (b) accept the null, reject the alternative hypothesis at the 95% confidence level

Question 4

Fit a simple linear regression of `salary` (dependent variable) against `salbegin` (independent variable).

I) What is the coefficient of `salbegin`?

II) What is the value of the residual standard error?

III) What is the t-value of the intercept term?

Question 5

Continuing with the linear model you formed in Question 4, we now wish to test the significance of the coefficients.

I) Determine the appropriate degrees of freedom for the test.

II) What is the value of the t-statistic for a hypothesis test with null that the coefficient of salbegin = 0.

III) What is the value of the t-statistic for a hypothesis test with null that the coefficient of salbegin = 2.

IV) Would we accept or reject the null hypothesis in part III at the 95% confidence level?
(a) Accept (b) Reject

Question 6

Look at the bank data set. There are 5 independent variables that might be important in determining the `salary`. Fit a new multiple regression model using these 5 variables.

I) What is the value of the residual standard error?

II) According to your model, how much more does a man make over a woman?

III) Do you think this model is a good fit?
(a) Yes (b) No