Name: ________________

Directions: MAKE SURE TO COPY YOUR ANSWERS TO A SEPARATE SHEET FOR SENDING ME AN ELECTRONIC COPY LATER.

Important note: Remember that in problems calling for R code, you are allowed to use any built-in R function, e.g. `choose()`, `sum()`, `integrate()` etc.

1. (15) Consider the marble example, Section 11.5. Find \( \text{Var}(Y | B = 2) \).

2. (15) Suppose in Equation (8.22) I wish to form an 88% confidence interval, instead of a 95% one. Give an expression, which must involve a call to one of the R functions we’ve used, to calculate the number I’ll use instead of 1.96.

3. Consider the “new, improved light bulbs” example in Section 9.6.2. **Note: Each of the parts here is independent of the others.**

   (a) (10) If we wished to have significance level \( \alpha = 0.10 \), sampling 50 bulbs, what should be our threshold for rejection, like the \( w \) in the example?

   (b) (15) Suppose we have 15 people test batches of 10 light bulbs, each performing a significance test as in the example. Suppose also that actually \( H_0 \) is true. Find the probability that at least 3 of the people reject \( H_0 \).

   (c) (15) Suppose it turns out that \( \bar{X} = 1624.2 \). Find the p-value.

4. In the baseball data, Section 11.7, I wanted to run separate regression analyses for catchers and starting pitchers.

   (a) (15) I extracted the two subsets of my original data frame `players`, naming them `catch` and `pitch`. Give one line of R code that creates `catch`.

   (b) (15) I ran regressions of weight on height in the two groups, with these results:

   ```R
   > summary(lm(catch$Weight ~ catch$Height))
   Call:
   lm(formula = catch$Weight ~ catch$Height)
   Residuals:
   Min 1Q Median 3Q Max
   -31.505 -7.603 -1.603 8.495 31.789
   Coefficients:
   Estimate Std. Error t value Pr(>|t|)
   (Intercept) -79.4301 67.9087 -1.17 0.246
   catch$Height 3.9019 0.9335 4.18 7.91e-05 ***
   ---
   > summary(lm(pitch$Weight ~ pitch$Height))
   Call:
   lm(formula = pitch$Weight ~ pitch$Height)
   Residuals:
   Min 1Q Median 3Q Max
   -45.236 -15.236 -0.193 12.527 65.883
   Coefficients:
   Estimate Std. Error t value Pr(>|t|)
   (Intercept) -126.6988 44.4338 -2.851 0.00477 **
   pitch$Height 4.4407 0.5943 7.472 1.89e-12 ***
   ``

   Find an approximate 95% confidence interval for the difference (catchers minus pitchers) between the slopes for the `Height` variables for the two groups.
Solutions:

1. 

\[ 0^2 \cdot 0.036/0.090 + 1^2 \cdot 0.048/0.090 + 2^2 \cdot 0.006/0.090 - 0.667^2 \]

2. 

\[-qnorm(0.06)\]

3.a 

\[ qgamma(0.90, 50, 0.001) / 50 \]

3.b 

\[ 1 - pbinom(2, 15, 0.05) \]

3.c 

\[ 1 - pgamma(16242, 10, 0.001) \]

4.a 

\[ catch <- subset(players, Position == "Catcher") \]

or 

\[ catch <- players[players$Position == "Catcher",] \]

4.b 

\[ (3.9019 - 4.4407) \pm 1.96 \sqrt{0.9335^2 + 0.5943^2} \]