***Week 2 Quiz***

* Read the following scenario + choose the correct set of hypotheses for the scenario:
* Since 2008, chain restaurants in CA have been required to display calorie counts of each menu item. Prior to menus displaying calorie counts, average calorie intake of diners at a restaurant was 1100 calories. After calorie counts started to be displayed on menus, a nutritionist collected data on # of calories consumed at this restaurant from a random sample of diners. Do these data provide convincing evidence of a difference in the average calorie intake of a diners at this restaurant?
* **H0: μ = 1100 HA: μ != 1100**
* Which of the following is the correct definition of the p-value?
* **P(observed or more extreme sample statistic | H0 true)**
* One-sided alternative hypotheses are phrased in terms of:
* **< or >**
* A Type 2 error occurs when the null hypothesis is
* **not rejected when it is false (***when alternative is true***)**
* T/F: Decreasing the significance level (α) will increase the probability of making a Type 1 error.
* **False 🡪** *Type I error rate = alpha*
* A study suggests average college student spends 2 hours per week communicating w/ others online. You believe this is an underestimate + decide to collect your own sample for a hypothesis test of random 60 students from your dorm + find on average they spent 3.5 hours a week communicating w/ others online. Which of the following is the correct set of hypotheses for this scenario?
* **H0: μ = 2 HA: μ > 2**
* Which of the following is the correct definition of the p-value?
* **P(observed or more extreme sample statistic | H0 true)**
* Two-sided alternative hypotheses are phrased in terms of:
* **≠**

* You set up a 2-sided hypothesis test for a population mean μ w/ a null hypothesis H0: μ= 100 + a significance level of α = 0.05. The p-value calculated from the data = 0.12, hence you failed to reject the null. Suppose after your analysis was completed + published, an expert informed you the true value of μ = 104. How would you describe the result of your analysis?
* **A Type 2 error was made b/c you failed to reject the null but μ is really not equal to 100.**
* A statistician is studying BP levels of Italians in ages 75-80. The following is some info about her study:
* Data were collected by responses to a survey conducted by email + no measures were taken to get info from those who did not respond to the initial survey email.
* The sample observations only make up about 4% of the population.
* The sample size is 2,047.
* The distribution of sample observations is skewed 🡪 easy to see, although not very extreme.
* The researcher is ready to use the CLT in the main part of her analysis.
* Which aspect of her study is *most likely* to prevent her from using the CLT?
* **(I), because the sample may not be random and hence observations may not be independent.**
* SAT scores are distributed w/ a mean = 1,500 + a SD = 300. You’re interested in estimating the average SAT score of 1st year students at your college. If you’d like to limit the margin of error of your 98% CI to 40 points, at least how many students should you sample?
* **306**

> sigma <- 300

> n <- 306

> SE <- sigma/sqrt(n)

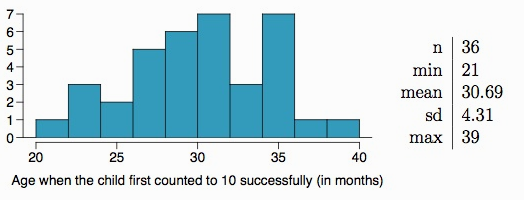
> (z.crit <- qnorm(1-(1 - .98)/2)) # ~ 2.32

[1] 2.326348

> (moe <- z.crit\*SE)

[1] 39.89654

* If it’s relatively riskier to reject the null when it might be true, should a smaller or a larger significance level be used?
* **smaller**
* Researchers investigating characteristics of gifted children collected data from schools in a large city on a random sample of 36 children IDed as gifted soon after the age of 4. The following histogram shows the distribution of ages (in months) at which these children first counted to 10 successfully. w/ some sample statistics.



* Suppose you read online that children first count to 10 successfully when 32 months old, on average. You perform a hypothesis test evaluating whether the average age at which gifted children first count to 10 is different than the general average of 32 months. What is the p-value of the hypothesis test? Choose the closest answer.
* **0.9656**

> n <- 36

> min <- 21

> x.bar <- 30.69

> trueNull.mu <- 32

> sigma <- 4.31

> max <- 39

>

> se <- sigma/sqrt(n)

> (z <- (x.bar - trueNull.mu)/se) # test statistic of .81

[1] -1.823666

> 1 - pnorm(x.bar,trueNull.mu,se)

[1] 0.9658987