Multiple Choice (3 pts each)

1) A 95% confidence interval for $\mu$ means that
   a. 95% of the confidence intervals generated will include the true mean $\mu$.
   b. 95% of the confidence intervals generated will include the sample mean $\bar{x}$.
   c. 95% of the possible population means $\mu$ will be included in the interval.
   d. 95% of the possible sample means $\bar{x}$ will be included in the interval.  

[3 pt]

2) Given that random variables $X_1, \ldots, X_{100}$ are independent and identically 
   distributed from $N(0,1)$ standard normal distribution, then $\bar{X}$ has distribution 
   a. approximately $N(0, 0.01)$
   b. exactly $N(0, 0.01)$
   c. exactly $N(0,1)$
   d. exactly $t_{100}$ (t-distribution with 100 df)

   we already have in our notes
   $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ 
   now, we see that
   $\bar{X} \overset{\text{approx}}{\sim} N(\mu, \frac{\sigma^2}{n})$ and only
   exactly $\bar{X} \overset{\text{exact}}{\sim} N(\mu, \frac{\sigma^2}{n})$

   when $X_1, \ldots, X_n \sim N(0,1)$

   [3pts]

3) For the $1-\alpha$ confidence interval for the population mean $\mu$ that has the form

   $\bar{X} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}$,

   the following assumptions are necessary:
   I. The data are i.i.d. normally distributed.
   II. The sample size is large.
   III. The population variance is unknown.

   a) I only.
   b) I and II only.
   c) II and III only.
   d) I, II, and III.

   [3pts]
4) Let \( \hat{\theta} \) be an estimator for the target parameter \( \theta \). Which of the following are random variables?

I. \( \hat{\theta} \)
II. a 90% CI for \( \theta \)
III. a pivot quantity for \( \theta \)
IV. \( \theta \)

a. I only
b. II only
(c. I and II only
(d. I, II and III

The z-score: \( z = \frac{X - \mu}{\sigma} \) is known as a "pivot quantity" as well as this is not a parameter! \( Z = \frac{\bar{X} - \mu}{\sigma / n} \)

At the key, here, is that our C.I. is random! The parameter is not... the estimate (X) is not! [3pts]

4) The phrase "regression to mediocrity" was coined by:
   a) William Goset
   b) Francis Galton
   c) Carl Friedrich Gauss
   d) Thomas Bayes

(All interesting contributors to Statistics)

(Did any of you read about the firing of Lawrence "Larry" Summers' position as President of Harvard University; as I suggested in class? --A very interesting interaction between science-statistics-sociology... I do not support his view, per se, I just find it extremely interesting, what occurred.)

5) For a large sample hypothesis test for the population mean

\( H_0 : \mu = 0 \)
\( H_a : \mu \neq 0 \)

if the standardized test statistic comes out to \( z = 1.8 \) then the p-value is

a) 0.0359
b) 0.051
c) 0.0719
d) 0.102

This is a two-sided test, so we need to find the area under both tails [3pts]

6) A sampling distribution is

Charlie Vollmer
March 9th, 2013

Quiz 3 Practice_Sp13.docx
Page 2 of 4
STAT307          Practice Quiz #3

a) always a normal distribution  
   not true... we have seen statistics that follow a t-distribution
b) the distribution for a random sample  
   No... but we need a random sample to be able to find the sampling dist. of our
c) the distribution for a statistic  
   this is just describing itself
   same distribution  
   [3pts]

d) the distribution for the r.v.'s  
   Y_1,...,Y_n when they are independent and from the

Write a short response to each of the following questions:

1) Define Type I and Type II errors.  
   go to notes... understand these well!
   Give an example where you would prefer making one type of error over the other
   (and explain why).
   well... would you prefer:
   1) locking up an innocent person
   2) letting a guilty woman walk free
   know which is which!

2) Define p-value.  
   go to notes... know this!
   Why is it often more practical for researchers to include p-values with the
   reject/not reject conclusions rather than test statistics?
   the p-value tells us "how certain" the researcher was at
   her conclusion... if \( \alpha = 0.05 \) and she finds p-value = 0.049, well...
   we know she rejects but just barely. If she tells us that
   she found a p-value = 0.0000031, then we know that she is
   definitely sure that rejecting the null is pertinent!

3) What is the advantage of a one-sided test over a two-sided test? Why wouldn't
   you always use a two-sided test?
   you can easily see that if we find a test statistic from our
   sample, that if it is two-sided the p-value is simply twice as
   large as the one-tail p-value.
   -But... if we have reason to believe that our test goes in a certain
   direction, then a one-tail p-value is smaller and justified and therefore our
   findings are more substantiated (we call this "more powerful")
4) What is a statistical reason for being skeptical of findings in medical journals reported with p-values of 0.05?

well... we actually expect (by construction) to find "significant" results 5% of the time. So, if you run this experiment just twenty times, we are almost surely guaranteed to find these "significant" results... we should demand higher standards!

5) What is the definition of a statistic?

a "statistic" is any function of our observed data which does not involve any unknown parameters

eq \bar{X} = \frac{1}{n} \sum x_i \text{ is a statistic}

\[ \frac{1}{n} \sum (x_i - \mu)^2 \] is not a statistic since it involves "\mu" which is our unknown parameter

(but... look \rightarrow \frac{1}{n-1} \sum (x_i - \bar{x})^2 does not involve any parameter!)