Practice Final Math 10 Spring 2006

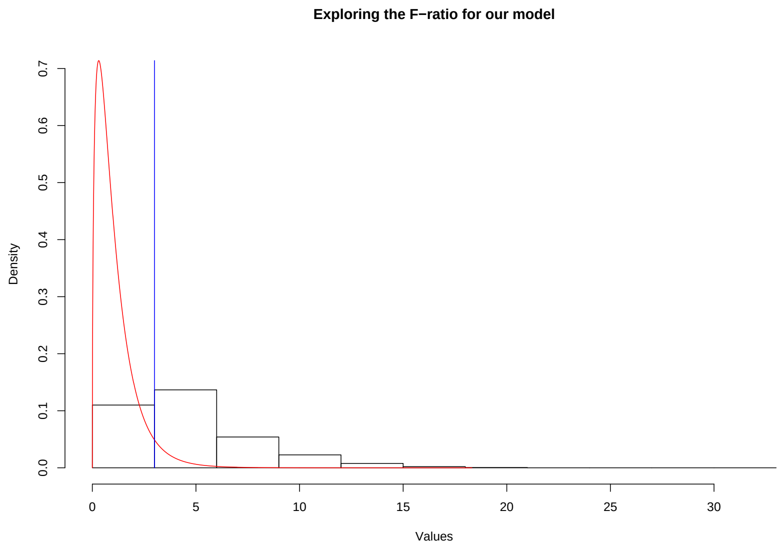
Practice Exam = Book Problems 1-8 described via Page #.Problem # as follows: 181.11, 345.27, 387.19, 501.23, 519.15, 543.13, 589.27, 619.11

and Non-Book Problems, 1-3 of the attached problems. Problem # 4 is the problem we did together on the last day of class.

1. Imagine that we are measuring the difference in the temperature of water which is initially at 180 degrees and left to sit in an insulated coffee cup for half an hour. We intend to compare this difference utilizing the mugs produced by one of 4 different mug manufacturers. We intend to test the mean difference for each mug by performing the experiment 8 times per mug. Based on price, we expect that the actual means should satisfy $\mu_1 > \mu_2 > \mu_3 > \mu_4$, and if the differences actually turn out to be something like $\mu_1 = 16$, $\mu_2 = 14$, $\mu_3 = 12$ and $\mu_4 = 10$ then we' d we'd like to be able to reject the *Null Hypothesis* that $\mu_1 = \mu_2 = \mu_3 = \mu_4$ with a high probability. To explore this we simulate this scenario 1000 times (assuming that each measurement of mug i is given by $N(\mu_i, 4)$). We compare a histogram of our simulated data with the curve corresponding to the distribution under the Null Hypothesis. We decide reject the null if the value of our statistic is greater than 3 which corresponds to a P-value of 0.05 (three is labeled via the vertical line in the picture).

   (a) What is the curve? (In other words, which distribution was plotted? and compute any relevant degrees of freedom.)

   (b) Can this plot be used to estimate the *power*? Explain. If it can, then estimate the *power*.



2. 40 students were invited by the administration to discuss "Campus Issues". We don't yet know the the class years of the administration's choices, but we intend to find out and then decide whether these years are uniformly distributed across the four possible class years (say 06, 07, 08 and 09). We intend to use the statistic

$$D = \sum \frac{|Obs - 10|}{\sqrt{10}}.$$

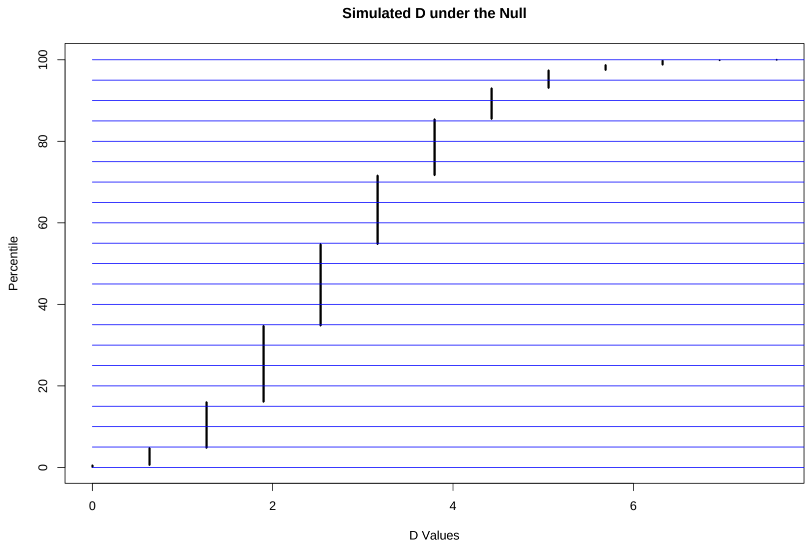
to test this hypothesis, but we are not sure how the statistic is distributed. So we simulate D's value under the hypothesis that the students were indeed randomly selected with regard to class year, and we notice our simulated values took on the possible values

$$\{0.00, 0.63, 1.26, 1.90, 2.53, 3.16, 3.79, 4.43, 5.06, 5.69, 6.32, 6.96, 7.59\}$$

and we graph the percentiles in the below figure. We also graphed a line at each five percent increment.

(a) Why will D be larger on average when the Null is false than when it is true?

(b) Would I reject the Null at the 95 percent confidence level if if when I run the experiment I find $D = 4.427$?

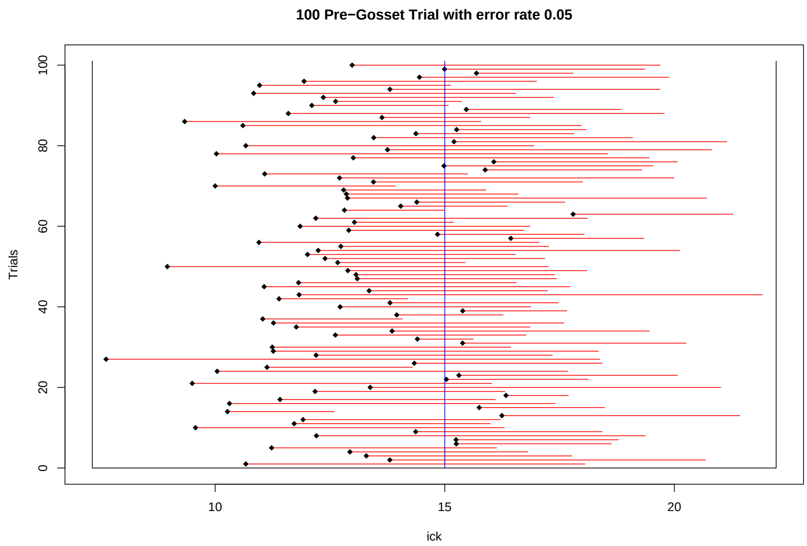


3. When brewing Guiness, a batch of the beer will have μ units of *ick* in it. Assume that when we test a sample from such a batch we measure $X = N(\mu, \sigma)$ units of *ick*, for some unknown σ. When all goes right in the brewing process $\mu = 15$, and when something goes wrong $\mu > 15$. Like Gosset, imagine we are going to test wether $\mu > 15$ for each batch using 4 samples and approximating a 90% confidence interval around our observed mean via $\hat{x} \pm (1.645)\frac{s}{2}$. Each batch that we determined had too much *ick* is then tested carefully by the Guiness lab and sometimes we are *wrong*.

    (a) When we are *wrong*, what type error have we made?

    (b) In the figure we see a graph of the first 100 confidence intervals that we measured which were determined by our test **or by the Lab** to have $\mu = 15$ units of *ick*. Is this what we had expected? and if not what happened?

    (c) From these intervals Estimate the 95 % confidence interval about the true error rate? What is this interval telling us?



4. In this problem we explore the book's rule of thumb that even for skewed distributions the t-test should work provided $N > 50$. To do this, we compute

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

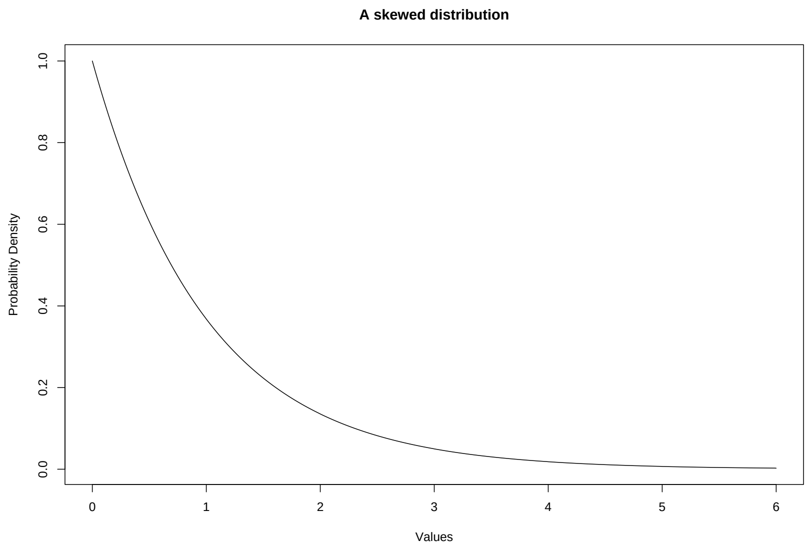
by independently constructing $n = 51$ samples using the skewed distribution described by the first of the below graphs. We then simulated the production of this t-value 10,000 independent times. A histogram of this simulated data is given in the second figure below, and this histogram is compared to the t-distribution with 50 degrees of freedom. In the last graph, we see the percentiles of our simulated data and also a line at each five percent increment.

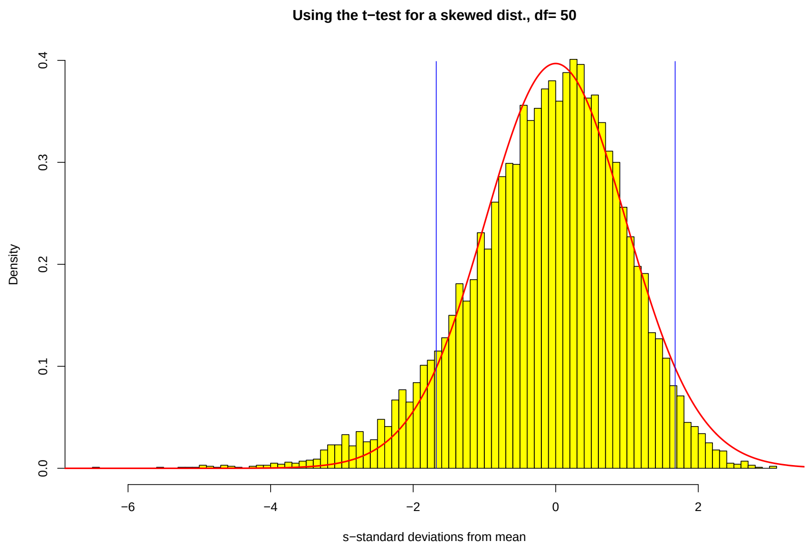
   (a) Assuming that the rule of thumb is correct, which t-value should have roughly 5 percent of the data smaller than it, and which t-value should have roughly 5 percent of the data larger than it. Call them t_{small} and t_{big}. (These represented in the plots as vertical lines.)

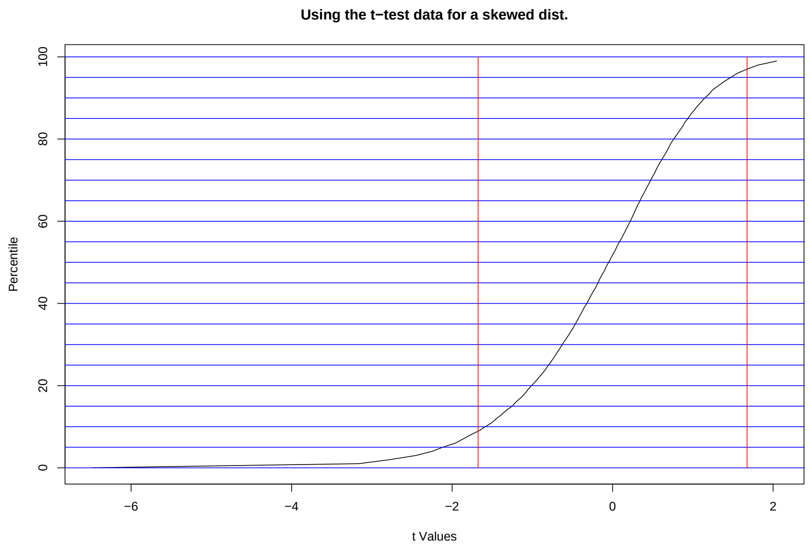
   (b) Estimate the actual percentage of our data that is less than t_{small}.

   (c) Estimate the probability that we would have seen the actual percentage estimated in part (b) due to chance error if the rule of thumb really worked.

   (d) (EC) Why is our histogram skewed to the left? (Notice, the original distribution was skewed to the right.)




Using the t−test for a skewed dist., df= 50
Density
0.0
0.1
0.2
0.3
0.4
−6
−4
−2
0
2
s−standard deviations from mean


Using the t−test data for a skewed dist.
Percentile
0
20
40
60
80
100
−6
−4
−2
0
2
t Values

Help Sheet:

Tables in back of book (if needed) and...

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

$$p^* = p_0 + \frac{\sqrt{p_0 q_0}(p_1 - p_0)}{(\sqrt{p_0 q_0} + \sqrt{p_1 q_1})}$$

Geometric: $P(X = x) = q^{x-1}p$, $E(X) = 1/p$, $Var(X) = q/p^2$.

Poisson: $P(X = x) = \frac{\lambda^x}{x!}e^{-\lambda}$, $E(X) = \lambda$, $Var(X) = \lambda$.

Binomial: $P(X = x) = \binom{n}{x}p^x q^{n-x} = \frac{n!}{(n-x)!x!}p^x q^{n-x}$.

$s^2_{pooled} = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{(n_1-1)+(n_2-1)}$

$MS_E = \frac{\sum e^2}{N-k}$

$MS_T = \frac{\sum(\bar{x}-\bar{\bar{x}})^2}{k-1}$

$r = \frac{\sum z_x z_y}{n-1}$

$b_1 = \frac{s_y}{s_x}r$

$b_0 = \bar{y} - b_1\bar{x}$

$s_e^2 = \frac{\sum(y-\hat{y})^2}{n-2}$

$SE(b_1) = \frac{s_e}{s_x\sqrt{n-1}}$

$SE^2(\hat{\mu}_\nu) = (x_\nu - \bar{x})^2 SE^2(b_1) + s_e^2/n$

$SE^2(\hat{y}_\nu) = SE^2(\hat{\mu}_\nu) + s_e^2$

$b_1 \pm t^\star_{n-2}SE(b_1)$

$\hat{y}_\nu \pm t^\star_{n-2} \times SE$