Statistics 515   Practice Exam 2 (Based on the Fall 2001 2nd Exam)

Part I: Answer seven of the following eight questions. If you complete more than seven, I will grade only the first seven. Five points each.

1) A fair coin is flipped 200 times. Use the central limit theorem to estimate the probability that at least 105 heads will be observed.

2) Define what is meant by "Type I error".

3) Define what is meant by "the p-value (or observed significance level) of a test".

4) A variety of methods can be used to check if a sample came from a population that follows a normal distribution. Which type of graphical display was designed specifically to help address this problem?

5) (Circle the correct answers) If you increase the sample size, you will be **more / less** likely to reject a false H_0 in a test of hypotheses.   If you increase α, you will be **more / less** likely to reject a false H_0 in a test of hypotheses.

6) A sample of size 20 results in $\bar{x} = 10.0$ and $s = 2.4$. Assume that the necessary assumptions are met and construct a 95% CI for μ.

7) (Circle the correct answers)   The hypothesis test and confidence intervals about a single variance are conducted using the **Z / t / χ^2 / F** distribution. The hypothesis test to compare two variances is conducted using the **Z / t / χ^2 / F** distribution.   In both cases the test is **very robust / not very robust**.

8) The output at the bottom of the page was generated using PROC TTEST on SAS for random samples from two data sets N and S.   Both samples appeared to have been drawn from populations that were approximately normally distributed.

(Circle the correct answers) At α=0.10 we **accept/reject** the null hypothesis that $\mu_N=\mu_S$. At α=0.01 we **accept/reject** the null hypothesis that $\mu_N=\mu_S$.

```
                        The TTEST Procedure

                           Statistics

                   Lower CL          Upper CL  Lower CL           Upper CL
Variable  group    N     Mean  Mean      Mean  Std Dev  Std Dev   Std Dev  Std Err
value     n       10   72.197  77.5    82.803   5.0985   7.4125    13.532    2.344
value     s       12   68.303  72.333  76.364   4.4938   6.3437    10.771   1.8313
value     Diff (1-2)   -0.947  5.1667  11.281   5.2371   6.8453    9.8851    2.931

                           T-Tests

          Variable    Method          Variances    DF    t Value   Pr > |t|
          value       Pooled          Equal         20      1.76     0.0932
          value       Satterthwaite   Unequal     17.9      1.74     0.0996

                       Equality of Variances

          Variable    Method     Num DF    Den DF    F Value    Pr > F
          value       Folded F        9        11       1.37    0.6165
```

Part II: Answer every part of the next two problems. Read each problem carefully, and show your work for full credit. Twenty points each.

1) A produce broker will deal with apples from an orchard only if he is quite certain that they are larger than 2.5 inches in diameter on average. To test this he randomly takes a sample of size 12 from the orchard's stock.

A) State the appropriate null and alternate hypotheses for testing if the broker will deal with the orchard's apples. Be sure to identify what the parameter(s) you are using mean in terms of the problem (e.g. if you use $\mu, p, \sigma^2, s^2, \bar{x}, \hat{p}$ say what the symbol stands for.)

B) A sample of size 12 is acquired and it has a mean of 2.758 and a standard deviation of 0.3942. Test the hypothesis in A at an $\alpha=0.01$ level and state what the broker should do in this case.

C) Besides the sample being randomly chosen, what other assumption(s) are required to trust the test in part B? If possible, check that the assumption(s) hold.

2) A candidate for political office wants to determine if there is a difference in his popularity between men and women. To test this he collects a sample of 250 men and 250 women and records how many of them plan on voting for him in the upcoming election.

A) State the appropriate null and alternate hypothesis for determining whether the candidate differs in popularity between men and women. Be sure to identify what the using mean in terms of the problem (e.g. if you use $\mu, p, \sigma^2, s^2, \bar{x}, \hat{p}$ say what parameter(s) you are the symbol stands for.)

B) Of those sampled, 105 of the men and 128 of the women plan on voting for the candidate. Report the p-value for the test of hypothesis in A.

C) Besides the sample being randomly chosen, what other assumption(s) are required to trust the test in part B? If possible, check that the assumption(s) hold.