

Name Jerry

Student Number _____

STA 442/2101 f2013 Quiz 4

Recall the test of difference between *variances* for the ~~Verbal~~ SAT data. Let X_i denote performance on the Verbal SAT and Y_i denote performance on the Math SAT for student i . For simplicity, let the sample variances $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ have n in the denominator rather than $n - 1$.

1. (1 point) Write down formulas for $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ in a form that shows how they are continuous functions of a collection of sample means.

$$\hat{\sigma}_1^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2, \quad \hat{\sigma}_2^2 = \frac{1}{n} \sum_{i=1}^n Y_i^2 - \bar{Y}^2$$

2. (3 points) We want a data vector \mathbf{D}_i to which we can apply the Central Limit Theorem. Then, we would write $\hat{\sigma}_1^2 - \hat{\sigma}_2^2 = g(\mathbf{D}_n)$, and the delta method would establish asymptotic normality. Show you know what's going on by writing down the data vector \mathbf{D}_i .

$$\mathbf{D}_i = \begin{pmatrix} X_i \\ X_i^2 \\ Y_i \\ Y_i^2 \end{pmatrix}$$

3. (3 points) You calculated the W_n statistic for this problem using R. Please

- Write the computed values of W_n and the p -value (two numbers) in the space below.

Every answer will be a little different because of the bootstrap.

Here are 3 answers $W_n = 2.19, p = 0.139$

$W_n = 2.01, p = 0.156$

$W_n = 2.06, p = 0.151$

- **Attach your R printout** – for this problem only. You cannot get any marks for Question 3 or Question 4 unless you attach the printout.
- **Circle** the value of W_n and the p -value on your printout.

4. (3 points) In plain, mostly non-statistical language, what do you conclude from this analysis? The answer is something about the variances of Verbal and Math SAT.

*There is not enough evidence to conclude that the variances are different, or
These results are consistent with no difference between variances.*

Printout for Quiz 4

```
> sat =
read.table("http://www.utstat.toronto.edu/~brunner/appliedf13/code_n_data/h
w/sat.data")
source("http://www.utstat.utoronto.ca/~brunner/appliedf13/code_n_data/lectu
re/Wtest.txt")
> sigmahat = var(sat[1:2]); sigmahat
      VERBAL      MATH
VERBAL 5359.686 1333.970
MATH   1333.970 4401.939
> set.seed(9999)
> n = dim(sat)[1]; n
[1] 200
> B = 1000; vstar = NULL
> for(i in 1:B)
+ {
+ choices = sample(1:n,n,replace=T)
+ datta = sat[choices,1:2]
+ vstar = rbind(vstar,diag(var(datta)))
+ }
> avarvar = var(vstar); avarvar
      VERBAL      MATH
VERBAL 265666.5 16214.5
MATH   16214.5 185584.5
> Wtest(L=cbind(1,-1),Tn=diag(sigmahat),Vn=avarvar)
      W      df      p-value
2.1901419 1.0000000 0.1388965
>
```