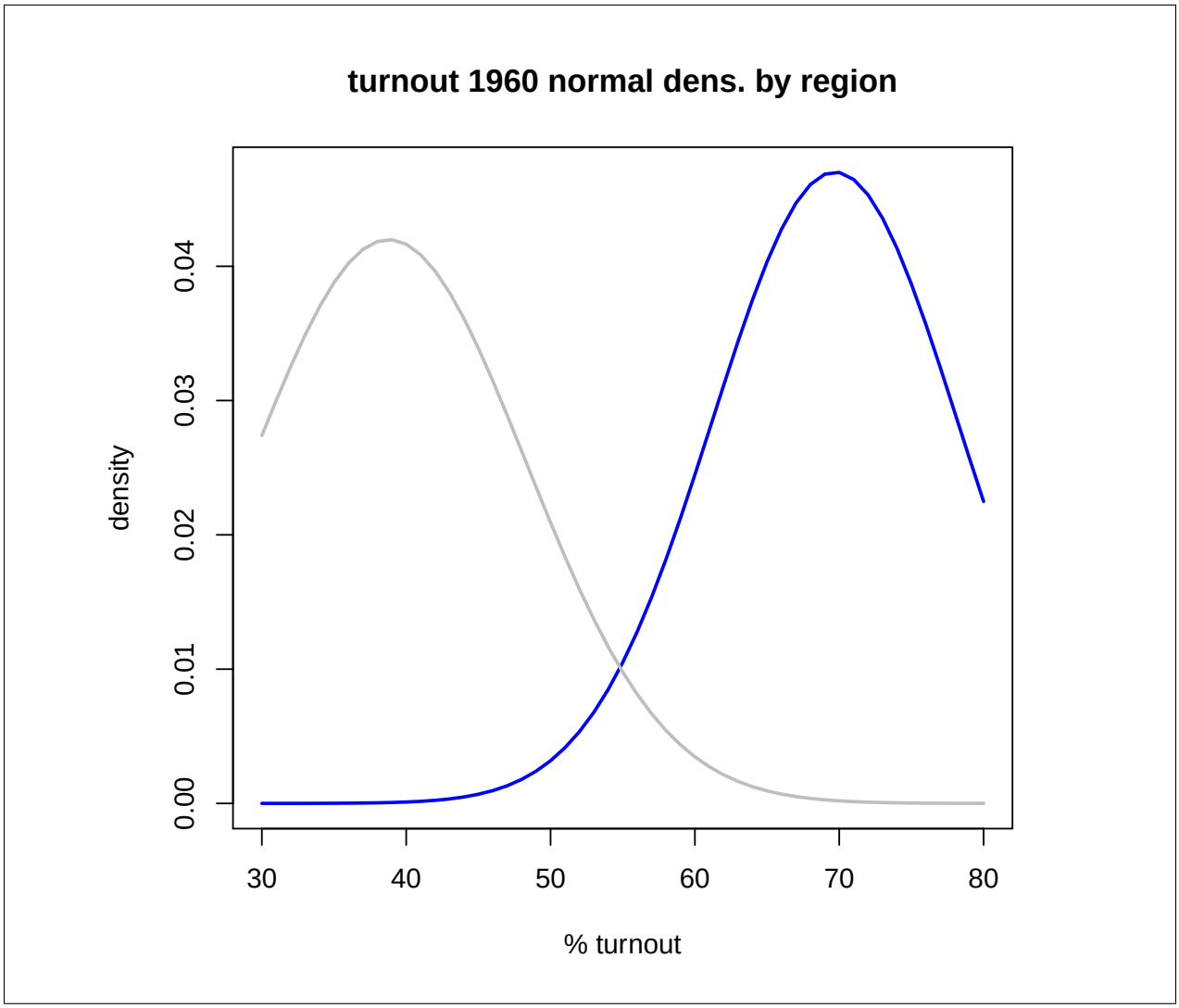
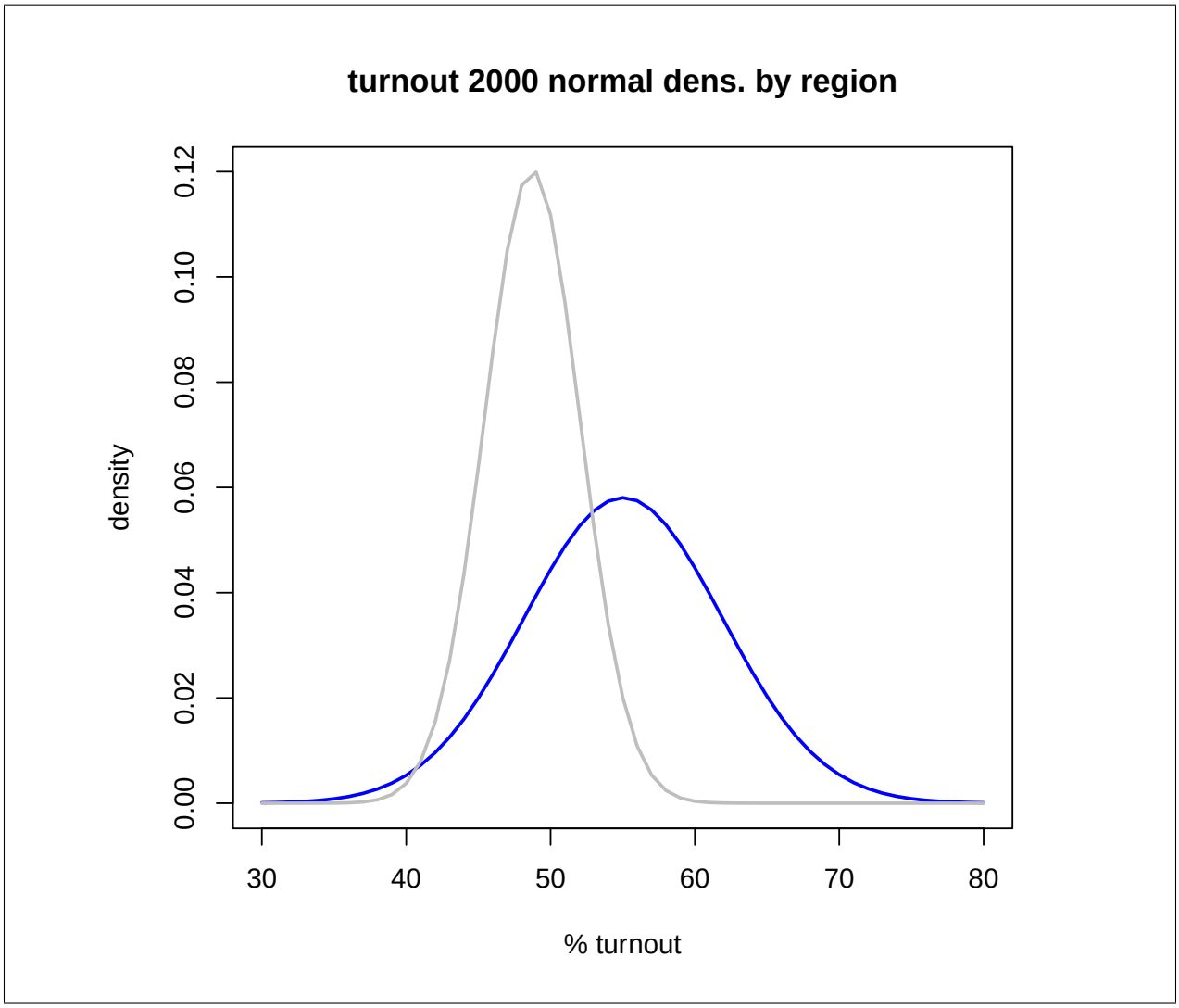
Exercise 1: 20 points

Students should have plot fairly similar to the below. Dock for deviations such as not labeled axes, no title, truncation of plot, etc.





Exercise 2: 20 points

1960: definitely not. 2000: possibly... there's a lot of overlap.

Exercise 3: 20 points

```
[1] "t-test 1960, delta=0"

$pval
[1] 1.731499e-10

$T
[1] 9.772513

$nu
[1] 27.77927
```

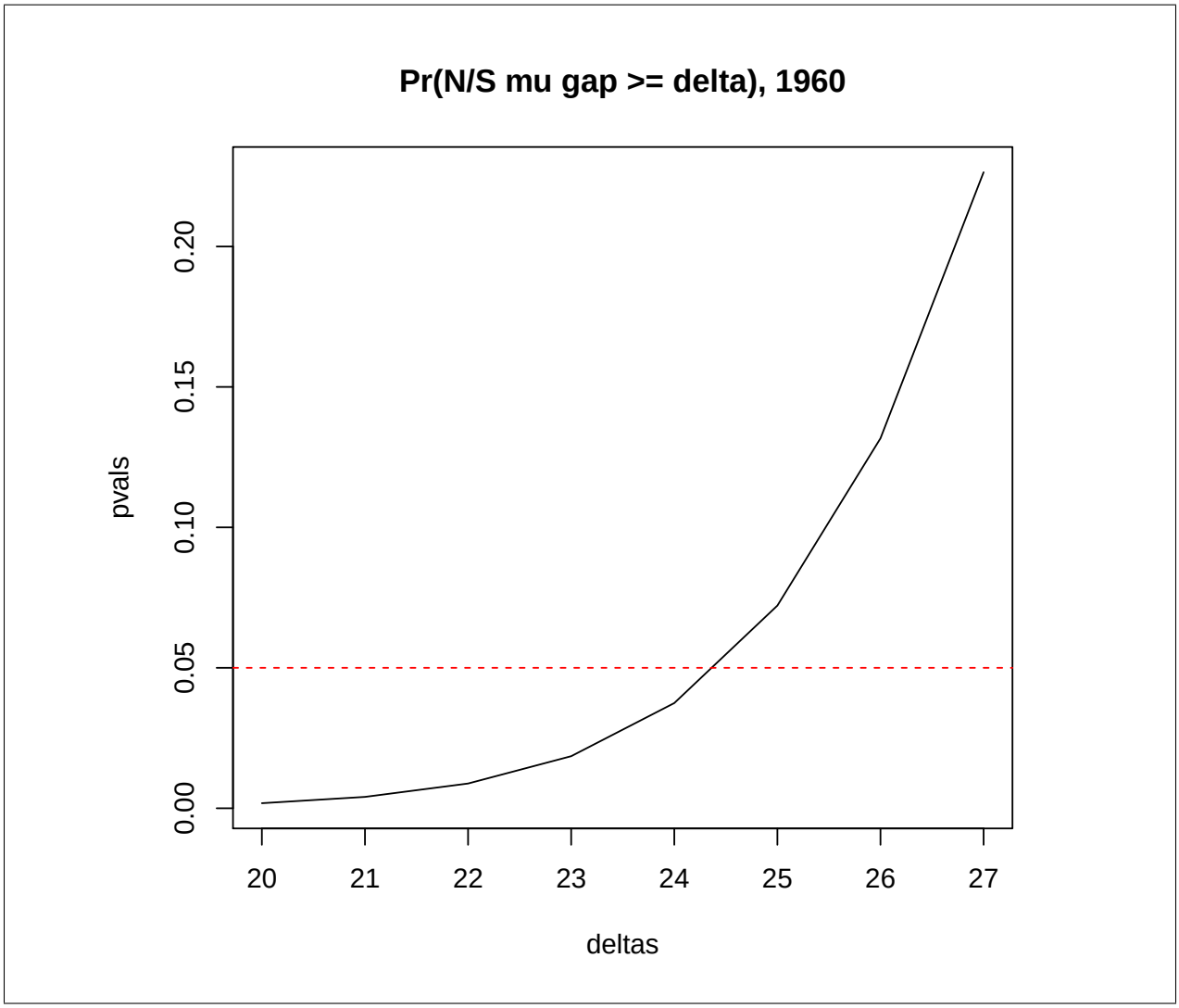
```
[1] "t-test 2000, delta=0"

$pval
[1] 8.906888e-05

$T
[1] 4.271799

$nu
[1] 48.93021
```

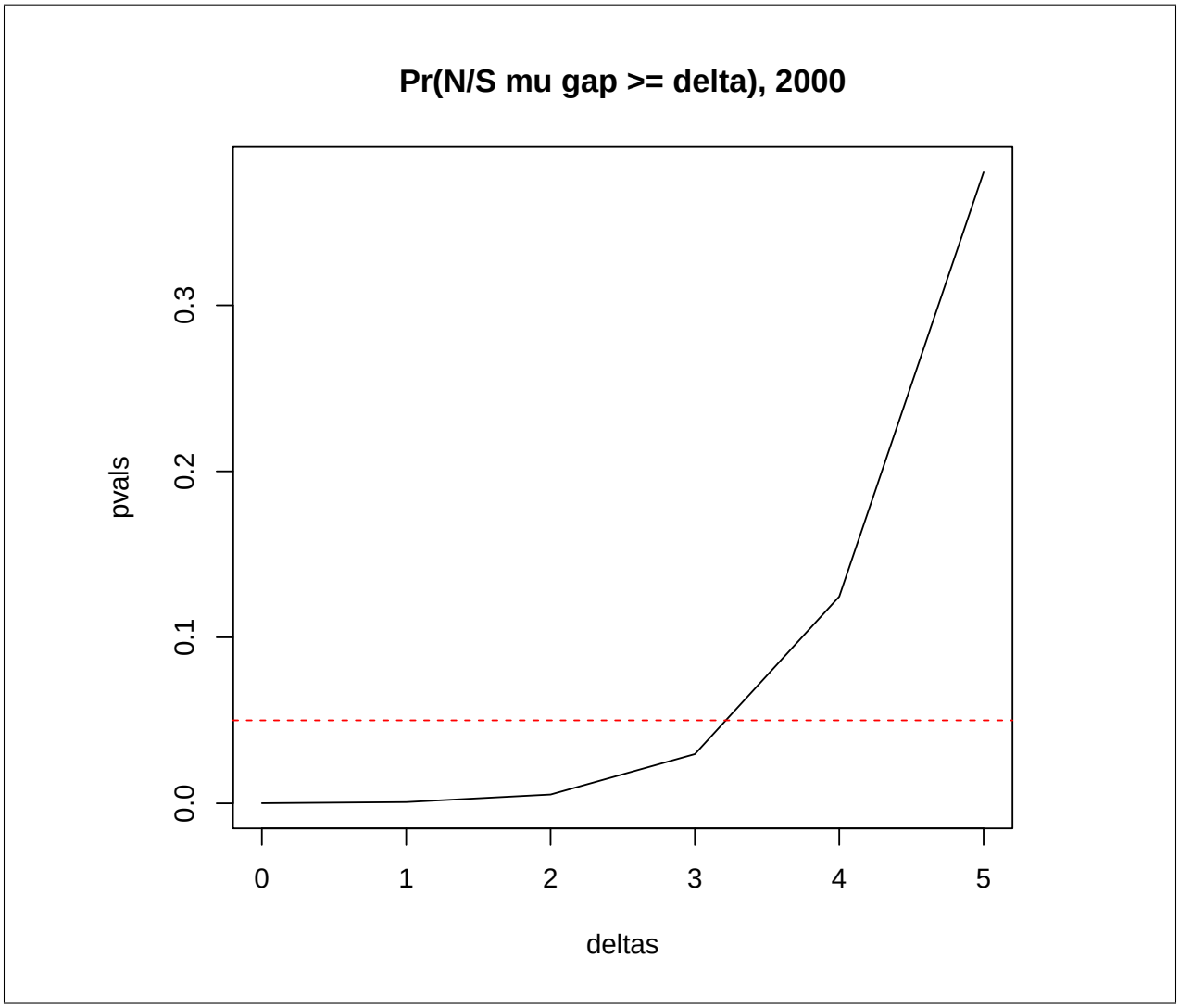
Students only need to show p-value. A correct interpretation is that the p-vals less than 0.05 allow us to reject the null hypothesis that the samples are drawn from populations with the exact same mean. They must also say that this is NOT interesting because even small differences in population means could lead to statistically significant results on this test ('interesting' is an OK substitute for 'statistically significant').

Exercise 4: 20 points

1960: $\Delta \approx 24.5$ is the maximum distance between the population means of South and North that is statistically significant. 2000: $\Delta \approx 3.25$

See below plots for reasining.


Pr(N/S mu gap >= delta), 1960
pvals
0.00
0.05
0.10
0.15
0.20
20
21
22
23
24
25
26
27
deltas



Exercise 5: 20 points

Observation of decreasing gap is supported since the maximum gap between parent populations in 1960 is very much smaller than that in 2000. If repeated for interceding years, we'd probably see Δ_{max} from exercise 4 decrease across years. Additionally, grade based on whether there is anything incorrect stated, and whether word limits (250-750) are met.