

22S:30/105, Statistical Methods and Computing
 Instructor: Cowles; Spring 2006, Homework 7

Due: Fri., 03/10 in class

- Textbook problems: 10.20, 10.22, 10.27, 13.16, 13.18, 13.20, 13.21
 - For problem 13.16, if you wish, you may use SAS for the stemplot and to compute \bar{x} . However, you will need to calculate the confidence interval by hand, because SAS does not make the assumption that σ is known. In addition to the 95% c.i. requested in part (b), calculate an 80% c.i.
- Refer to textbook problem 13.17. The data is in a file called “crank.dat” on the course web page.
 - Use SAS to make the stemplot in part (a). Write a brief description of the shape of the distribution of values.
 - Do part (b) by hand. Assume that σ is known as stated in the problem.
 - Use SAS proc means to compute a 95% c.i. for the mean μ . Proc means does *not* assume that σ is known. Is the 95% c.i. that it computes wider or narrower than yours from part (b)?
- Download the dataset entitled ‘pprv.dat’ from the course web page. Use SAS to produce the following summary table giving the number of observations for each smoking group.

Refer to the Lab 3 handout for how to use formats, labels, and proc freq so as to produce the table exactly as it looks below. Turn in your SAS code and the printed table.

The FREQ Procedure

Smoking History

| smokhist | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|----------|-----------|---------|-------------------------|-----------------------|
| Current | 7 | 9.72 | 7 | 9.72 |
| Former | 44 | 61.11 | 51 | 70.83 |
| Never | 21 | 29.17 | 72 | 100.00 |

- OPTIONAL EXTRA CREDIT PROBLEM: This problem concerns the uniform distribution, described in the textbook on p. 235.

- Use the following SAS code to generate 1000 values from the uniform distribution. (This SAS code is posted as “uniform.sas” under “Datasets” on the course web page and can be copied and pasted from there into the SAS program editor window.)

```
data uni;
seed = 32542 ;
retain seed ;
do i = 1 to 1000 ;
  y = ranuni(seed) ;
  output ;
end ;
drop seed ;
run ;
```

- Use SAS to identify a simple random sample of size 30 from the dataset you created in part (a).
- Use SAS to compute the sample mean \bar{x} and the sample standard deviation s from your sample.
- For a population of values from a theoretical uniform distribution, the population mean μ is 0.5 and the population standard deviation σ is $\sqrt{\frac{1}{12}}$.
 - What are the mean and standard deviation of the sampling distribution of \bar{x} s for samples of size 30 drawn from this population?
 - In this case, the sample size $n = 30$ is large enough that the sampling distribution is approximately normal (according to the Central Limit Theorem). What is the probability of getting an \bar{x} as large as, or larger than, the one you actually got?