Midterm 2 Review — Stat 100

For the midterm: YOU WILL NEED a calculator.

Midterm 2 will cover 5.6 through 8.3, excepting 5.7, 6.6, and 6.7. Most problems will be word problems (like the homework) you must decode to use the ideas.

One question will be a word problem involving a binomial distribution as in 5.6, for which you will use the binomial distribution tables. On the sample exams, this type of question is mostly on midterm 1 exams.

CHAPTER 6

For the exam, you will be given a copy of the table for the standard normal distribution probabilities.
Given a random variable Z with a standard normal distribution, you must be able to use the tables to compute probabilities like Prob$(Z < -1.51)$, Prob$(|Z| < 1.51)$, Prob$(Z > 1.51)$.
Also be able to solve for numbers giving certain probabilities, for example, find c satisfying Prob$(Z < c) = .04$, or Prob$(|Z| < c) = .92$, or Prob$(Z > c) = .04$, or Prob$(3c < Z < 1.5) = .5$.

For a random variable X, you MUST know how to standardize. If X has mean μ and standard deviation σ, then $Z = (X - \mu)/\sigma$ has mean 0 and standard deviation 1. If X has a normal distribution, then so does its standardized version Z. Be able to do all the problems of the sort described in the first paragraph for any normal r.v., by converting the problem to a problem for the standardized r.v.

Suppose X is an r.v. with distribution Binomial(n, p).
(For example: X is the number of heads seen in n independent flips of a coin which shows heads with probability p.)
Then X has mean $\mu = np$ (common sense) and standard deviation $\sigma = \sqrt{np(1-p)}$.
Let Z be the standardized version of X. Given p, as n gets large the distribution of Z approaches the standard normal distribution. Be able to work word problems (as in 6.5) in which you recognize an r.v. is binomial, and compute a probability by using the normal approximation. Stat 100 rule of thumb: this approximation is justified/reasonable if np and $n(1-p)$ are both at least 15.
When you make the normal approximation to a binomial r.v. X, use the Continuity Correction: for example, for an integer n, when standardizing to compute Prob$(X \leq n)$, first replace this with Prob$(X \leq n + 1/2)$.

CHAPTER 8

We consider the population mean.
There are two kinds of problems: for some number α,
I. find a $100(1-\alpha)\%$ confidence interval or $100(1-\alpha)\%$ standard error, or
II. determine how large a sample must be in order that the $100(1-\alpha)\%$ standard error be smaller than some given number.

On the midterm, you will see one or both of these kinds of problems, in a word problem you must decode. Answers are based on supposing $\bar{X}$ is normal, and there are two possible justifications for this (see CHAPTER 7).

In each case you have to be able to find the number "$z_{\alpha/2}$". This number is defined by the requirement that for Z standard normal, $\text{Prob}(Z > z_{\alpha/2} = \alpha/2)$.
The meaning of the number is that $\text{Prob}(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = 1-\alpha$.
The number $z_{\alpha/2}$ is $-c$ where c is the number which satisfies $\text{Prob}(Z \leq c) = \alpha/2$.
You can find c from the normal tables. Just practice this for a few values.
(For example: 92% confidence interval involves $\alpha/2 = .04$ and $z_{\alpha/2} = 1.75$.)

Type I.
The $100(1-\alpha)\%$ standard error (for the estimate of μ by $\bar{X}$) is $b = z_{\alpha/2}(\sigma/\sqrt{n})$.
If σ is unknown and n is large, then for an estimate you can use $b = z_{\alpha/2}(s/\sqrt{n})$, where s is the sample standard deviation.
The $100(1-\alpha)\%$ confidence interval for μ is the interval $(\bar{X}-b, \bar{X}+b)$.

For example, if $100(1-\alpha)\% = 95\%$, then the meaning of the confidence interval is that in 95% of the random samples, the true mean μ will fall in the confidence interval.

Type II.
If you require a sample size n big enough that the $100(1-\alpha)\%$ standard error is less than some given number d, then what you are requiring is

$$z_{\alpha/2}(\sigma/\sqrt{n}) \leq d.$$

So, solve n in this condition, and if n is not an integer then round it up to the nearest integer.

CHAPTER 7

Given some random variable X (with mean μ and standard deviation σ) and random samples of size n, you have the random variable $\overline{X} := (X_1 + \cdots + X_n)/n$. Under the random sample assumption, $\overline{X}$ has mean μ and standard deviation $\sigma/\sqrt{n}$.

You will have to be able to compute or estimate Prob$(c < \bar{X} < d)$ for numbers c, d. You can do this by knowing or estimating that the distribution of $\bar{X}$ is normal. Here the *estimate* is justified by the Central Limit Theorem for n large (Stat 100 rule of thumb: $n > 30$). Also, for ANY n (especially: small n), if X is normal then so is $\bar{X}$.

The handouts on the Central Limit Theorem and Consistency and Bias summarize some key ideas (the handouts on Law of Large Numbers and the Normal Distributions may also help). The Central Limit Theorem will be applied, explicitly or implicitly, in word problems.

SUMMARY

On the midterm expect word problems involving calculations from the following list.
1. binomial distribution (5.6)
2. normal approximation to binomial
3. central limit theorem
4. find confidence interval (8.3)
5. find sample size (8.3)

There will probably also be some problem to test understanding of items such as the Central Limit Theorem, Consistency and Bias, sampling distribution, and the meaning of a confidence interval. The format might be True/False questions, but you might also be asked to describe something (such as the meaning of a 95% confidence interval) in your own words.