The second test (Test #2) will occur on Monday, November 13, 2006 and will cover the following sections of the text: §§4.1–4.2, §§5.1–5.4. Below are my comments on each of these sections.

It goes without saying—but not without writing—that it is assume that you have a good go at the assigned exercises.

Chapter 4: Discrete Probability Distributions

§4.1 Probability Distributions. The notion of a random variable is introduced, there are two types, discrete and continuous.

Discrete random variables are characterized by their probability distribution function, \( P(x) \). If \( x \) is a discrete r.v., then its pdf is defined by

\[
P(a) = P(x = a)
\]

The pdf can be displayed in a table, a probability histogram, or a mathematical function. The values of the pdf are numbers between 0 and 1, and \( \sum P(x) = 1 \). There are several exercises that query you on these points.

The mean and variance of a discrete r.v. are computed as follows:

\[
\mu = \sum x \cdot P(x) \tag{1}
\]

\[
\sigma^2 = \sum (x - \mu)^2 \cdot P(x) = \sum x^2 \cdot P(x) - \mu^2 \tag{2}
\]

The boxed expressions are the ones used for calculation, as demonstrated in class. If the distribution is given in a table, one has \( x \cdot P(x) \) and \( x^2 \cdot P(x) \) columns. The column totals of these two columns are \( \sum x \cdot P(x) \) and \( \sum x^2 \cdot P(x) \), respectively. These, in turn, are used in the two formulas (1) and (2).

You should do the assigned exercises, be able to compute some simple probabilities, and to compute the mean, variance and standard deviation.

§4.2 Binomial Distributions. The binomial random variable plays an important role in statistics.

Review the definition of a binomial experiment; standard notations are \( n, p \) and \( q \).

Recall that a binomial random variable counts the number of successes in the binomial experiment. You should be able to read a experimental situation, and “parse” it by identifying the underlying random variable, identifying \( n, p \) and \( q \).

The binomial pdf is given by an algebraic formula, however, we’ll be computing binomial probabilities using the binomial tables, or by the infamous, all-knowing calculator.

In some of the examples, I increased computing efficiency by using the complementary rule, for example,

\[
P(x > 5) = 1 - P(x \leq 5)
\]

Other probability properties can be exploited as well, such as the addition rule:

\[
P(x < 3 \text{ or } x > 5) = P(x < 3) + P(x > 5)
\]

The mean and variance of a binomial random variable are given by algebraic formulas:

\[
\mu = np \quad \sigma^2 = npq \quad \sigma = \sqrt{npq}
\]
Chapter 5: Normal Probability Distributions

§5.1 Introduction: Standard Normal Distributions. A general introduction to the normal probability distribution curve and its geometric properties.

The major topic in this section is the standard normal random variable, usually denoted by $z$, and its distribution, the standard normal distribution ($\mu = 0$ and $\sigma = 1$).

In this section we learned how to compute the area under a standard normal curve using the standard normal tables and the infamous, all-knowing calculator. Using area arithmetic and the fact that the total area is 1 square unit, we can add and subtract areas (obtained from the tables or from the infamous, all-knowing calculator) to compute the area we want.

Be prepared to compute various areas under the standard normal curve. Do the assigned exercises to prepare yourself.

§5.2 Normal Distributions: Finding Probabilities. This section continues §5.1, but for an arbitrary normal distribution. The key is to compute the $z$-score of the endpoints of the event. Recall,

$$P(a \leq x \leq b) = P(z_a \leq z \leq z_b)$$

where, $z_a$ and $z_b$ are the $z$-scores of $a$ and $b$, respectively; that is,

$$z_a = \frac{a - \mu}{\sigma} \quad z_b = \frac{b - \mu}{\sigma}$$

To compute the probability given in the left-hand side of equation (3), we compute the probability given in the right-hand side of equation (3).

You should do all the exercises that I assigned. Be prepared for these problem types.

§5.3 Normal Distributions: Finding Values. In §§5.1–5.2 we looked up area/probabilities using a given $z$-value ($z$-score, if you prefer). In this section, we use the tables in the opposite way, given an area or probability, we look up the corresponding $z$-value ($z$-score).

Basically, once the $z$-value is found, we then compute the corresponding $x$-value:

$$x = \mu + z \cdot \sigma$$

Do the assigned problems and be prepared to do these problem types.

§5.4 Sampling Distributions and the CLT. This is a pivotal section the content is used throughout the rest of the course.

A sampling distribution is the probability distribution of a sample statistics. In this section the statistics of interest is the sample mean $\bar{x}$, based on a sample size of $n$.

**Properties of the Sampling Distribution of a sample mean.** The mean, variance, and standard deviation of $\bar{x}$ when sampling from a random variable with mean $\mu$ and variance $\sigma^2$ are

$$\mu_{\bar{x}} = \mu$$

$$\sigma^2_{\bar{x}} = \frac{\sigma^2}{n}$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$
Armed with $\mu_{\bar{x}}$ and $\sigma_{\bar{x}}$, the techniques of sections §§5.2–5.3 are used. For example to compute a $z$-score for a value of a sample mean $\bar{x}$ we use the formula:

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

(5)

When sampling from a normal population, $\bar{x}$ is also normally distributed as well.

**The Central Limit Theorem.** This is one of the most important results in probability and statistics! Basically, it says, when you are working with a sample mean, $\bar{x}$, and when the number of samples, $n$, is large, $n \geq 30$, the sampling distribution of $\bar{x}$ is approximately normally distributed.

As a consequence of the CLT, we can then use the standard normal tables to compute probabilities concerning $\bar{x}$.

Do the assigned exercises for this section, the only thing new is the computation of the standard deviation of the sample mean, $\sigma/\sqrt{n}$ instead of just $\sigma$. The use of the tables. Note that

$$P(a \leq \bar{x} \leq b) = P(z_a \leq z \leq z_b)$$

(6)

where, $z_a$ and $z_b$ are the $z$-scores of $a$ and $b$, respectively; that is,

$$z_a = \frac{a - \mu}{\sigma/\sqrt{n}} \quad z_b = \frac{b - \mu}{\sigma/\sqrt{n}}$$

(7)

Note that equations (6) and (7) are the same as equations (3) and (4) with one important exception, the computation in the denominator of the $z$-score; otherwise, all techniques of §§5.1–5.3 apply! Remember that!

Do the assigned exercises, and be prepared to do these problem types.

The test, no doubt, will have some short answers (true/false, fill in the blank) and some questions involving computation. Good luck, but more importantly, good knowledge. I shall attempt to construct fair test over these topics.

Regards, D P S