Lecture 6: **Chapter 6: Normal Probability Distributions**

A **normal distribution** is a continuous probability distribution for a random variable x. The graph of a normal distribution is called the normal curve. A normal distribution has the following properties.

1. The mean, median, and mode are equal.
2. The normal curve is bell shaped and is symmetric about the mean.
3. The total are under the normal curve is equal to one.
4. The normal curve approaches, but never touches, the x-axis as it extends farther and farther away from the mean.
5. Between $\mu - \sigma$ and $\mu + \sigma$ (in the center of curve) the graph curves downward. The graph curves upward to the left of $\mu - \sigma$ and to the right of $\mu + \sigma$. The points at which the curve changes from curving upward to curving downward are called **inflection points**.
6. The Empirical Rule: Approximately 68% of the area under the normal curve is between $\mu - \sigma$ and $\mu + \sigma$. Approximately 95% of the area under the normal curve is between $\mu - 2\sigma$ and $\mu + 2\sigma$. Approximately 99.7% of the area under the normal curve is between $\mu - 3\sigma$ and $\mu + 3\sigma$.

6-2 The standard Normal Distribution

Def: The **standard normal distribution** is a normal probability distribution that has a mean of 0 and a standard deviation of 1.

Finding Area under the Standard Normal curve using Table A-2:

When using table A-2, it is essential to understand the following points.

I) Table A-2 is designed only for the standard normal distribution, which has a mean of 0 and a standard deviation of 1.

II) Each value in the body of the table is the cumulative area from the left op to a vertical line above a specific value of *z*. Recall: the z-score allows us to transform a random variable x with mean μ and standard deviation σ into a random variable *z*. $z = \frac{x-\mu}{\sigma}$ with mean 0 and standard deviation of 1.

III) When this transformation takes place, the area that falls in the interval under the nonstandard normal curve is the same as that under the standard normal curve with the corresponding z-boundaries.

IV) Table A-2 is on two pages. One for positive z scores and one for negative z scores.

Example 1: Find the indicated area under the standard normal curve.

a) Between z=0 and z = 1.96
b) To the right of z = 1.64
c) To the left of z = 1.54
d) To the right of z = -0.95
e) To the left of z =-2.57
f) Between z=-0.44 and z = 1.66
g) Between z =1.66 and z = 2.97

Notation:

$P(a<z<b)$ denotes the probability that the z score is between a and b.

$P(z>a)$ denotes the probability that the z score is greater than a.

Note: $P(z>a) = 1 - p(z<a)$

$P(z<a)$ denotes the probability that the z-score is less than a.

$P(z=a) = 0$ the probability of getting any single exact value is 0.

Finding z scores with given probabilities:

Example 2: $P(0<z<b) = 0.4923$ $P(a<z<0) = 0.3925$ $P(z<a) = 0.3132$

Example 3: Find the z score so that the area to the left of the z score is 0.32.

Remember that if the area to the left of the z sore is less than 0.5, the z score must be less than 0. If the area to left of the z score is greater than 0.5, the z score must be greater than 0.

Example 4: Find the z score so that the area to the right of the z score is 0.4322.

Example 5: Find the z scores that divide the middle 90% of the area in the standard normal distribution from the area in the tails.

Example 6: Find the z score that divides the bottom 95% from the top 5%.

For z scores above 3.49, use 0.9999 as an approximation of the cumulative area from the left. For z scores below -3.49, use 0.0001 as an approximation of the cumulative area from the left.

6-3 Applications of Normal distributions:

The standard score, or z score, represents the number of standard deviation a random variable, x, falls from the mean, μ. To transform the random variable to a z score, use $z = \frac{x-\mu}{\sigma}$ (round to 2 decimal places).

Example 7: a survey was conducted to measure the height of American males. In the survey, respondents were grouped by age. In the 20-29 age group, the heights were normally distributed, with a mean of 69.2 inches and a standard deviation of 2.9 inches. A study participant is randomly selected.

a) Find the probability that his height is less than 66 inches.

b) Find the probability that his height is between 66 and 72 inches.

c) Find the probability that his height is more than 72 inches.

Transforming a z score to an x value: $x = \mu + z\sigma$

(Make z negative if it is to the left of the mean.)

Example 8: Scores for a civil service exam are normally distributed, with a mean of 75 and a standard deviation of 6.5. To be eligible for civil service employment, you must score in the top 5%. What is the lowest score you can earn and still be eligible for employment?

Example 9: The lengths of pregnancies are normally distributed with a mean of 268 days and a standard deviation of 15 days. If we stipulate that a baby is premature if the length of pregnancy is in the lowest 4%, find the length that separates premature babies form those who are not premature.

Example 10: a pediatrician obtains the heights of her 200 3-year-old female patients. The heights are approx. normally distributed, with mean 38.72 in and standard deviation of 3.17 in. Use the normal model to a) determine the proportion of the 3-year-old females that have a height less than 35 inches. b) Find the percentile rank of a 3-year-old female whose height is 43 inches. Hint: the k^{th} percentile dives the lower k% of a data set from the upper (100-k)%. c) Compute the probability that a randomly selected 3-year-old female is between 35 and 40 inches tall, inclusive. d) Find the height of a 3-year-old female at the 20^{th}percentile. E) Determine the heights that separate the middle 98% of the distribution.

6-4 Sampling Distributions and Estimators

In this section, we will study the relationship between a population mean and the means of samples taken from the population.

Def: The **sampling distribution of the mean** is the probability distribution of sample means, with all samples having the same sample size n.

Example 11: A population consists of the values 1, 2, 5.

a) Find the population mean.
b) List all of the possible samples (with replacement) of size n=2 along with the sample means and their individual probabilities.
c) Find the mean of the sampling distribution of means.
d) Do the sample means target the value of the population mean?

In general, the distribution of sample means will have a mean equal to the population mean. Sample means therefore tend to target the population mean.

Example 12: For the population of 1, 2, 5.

a) Find the population proportion of odd numbers.

b) List all of the possible samples (with replacement) of size n=2 along with the sample proportions of odd numbers and their individual probabilities.
c) Find the mean of the sampling distribution of proportions.
d) Do the mean of sampling distribution of proportions for odd numbers target the value of the population proportion for odd numbers?

In general, the sampling distribution of proportions will have a mean that is equal to the population proportion.

If some of sample statistics are good to target the population parameter, then we call them unbiased estimators.

Mean, variance, and proportion are unbiased estimators. The statistics that do not target population parameter are median, range, and standard deviation. The bias is relatively small in large samples for standard deviation, thus s is often used to estimate σ.

As you have noticed, for small samples we have used sampling with replacement; however sampling without replacement would avoid wasteful duplication. We are interested in sampling with replacements because a) if the sample size is relatively smaller than the large population, then there is no significant difference between sampling with replacement or without replacement. b) Sampling with replacement results in independent events, and independent events are easier to analyze and they result in simpler formulas.

Example 13: Try #9 on section 6-4

6-5 The Central Limit Theorem

1. **If samples of size n, where n>30, are drawn from any population with a mean μ and a standard deviation of σ, then the sampling distribution of sample means approximates a normal distribution. The greater the sample size, the better the approximation.**
2. **If the population itself is normally distributed, the sampling distribution of sample means is normally distributed for any sample size n. In either case,**

the sampling distribution of sample means has a mean equal to the population mean. $\mu_{\bar{x}}=\mu$

3. **And the sampling distribution of sample means has a variance equal to 1/n times the variance of the population and a standard deviation equal to the population standard deviation divided by the square root of n.**

The standard deviation of the sample means is often called the **standard error of the mean.**

Probability and the Central Limit Theorem

We can find the probability that a sample mean, $\bar{x}$, will fall in a given interval of the sampling distribution. To transform $\bar{x}$ to a z score, you can use the equation $z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$

Example 14: The mean rent of an apartment in a professionally managed apartment building is $780. You randomly select nine professionally managed apartments. What is the probability that the mean rent is less than $825? Assume that the rents are normally distributed, with a standard deviation of $150.

Example 15: Credit card balances are normally distributed, with a mean of $2870 and a standard deviation of $900. A) What is the probability that a randomly selected credit card holder has a credit card balance less than $2500? B) You randomly select 25 credit card holders. What is the probability that their mean credit card balance is less than $2500?

Example 16: During a certain week the mean price of gasoline in the New England region was $\mu = \$1.080$ per gallon. What is the probability that the mean price $\bar{x}$ for a sample of 32 randomly selected gas stations in that area was between $1.075 and $1.090 that week? Assume the standard deviation is $0.045

Example 17: Try it yourself #16 on section 6-5