

22S:30/105
Statistical Methods and
Computing

Normal Distributions

Lecture 4
January 30, 2009

Kate Cowles
374 SH, 335-0727
kcowles@stat.uiowa.edu

Using density curves to describe the
distribution of values of a quantita-
tive variable

- Imagine the heights of 100,000 men who completed physical exams as part of a national health survey.
- We might make relative frequency histograms of these height data using successively smaller-width intervals.

Density curve – a curve that describes the overall pattern of a distribution

- total area under a probability density curve is 1.0
- the curve never drops below the horizontal axis

Measures of center and spread can be used to describe density curves.

- To distinguish between these measures in the idealized curve vs. in actual sample data, we use different symbols:
 - μ for the *mean* of a density curve
 - σ for the *standard deviation* of a density curve

Some characteristics of the normal distribution

1. Normal distribution is *symmetric*
 - (a) The proportion of the values of a normal random variable that are less than $\mu - z \sigma$ is equal to the proportion of the values that are greater than $\mu + z \sigma$
 - (b) The proportion of the values of a normal random variable that are less than $\mu + z \sigma$ is equal to the proportion of the values that are greater than $\mu - z \sigma$
 - (c) The mean is equal to the median

Normal distributions

- characterized by a symmetric, smooth bell shape
- also called “Gaussian distributions” (after Karl Gauss)
- The normal distribution is a *mathematical model* that provides a good representation of the values of many kinds of real quantitative variables
 - Analogy: No room is *perfectly* rectangular in shape, but the geometric model of a rectangle is good enough to enable you to measure the room and buy the right amount of carpet!

2. There are lots of different normal distributions, defined by different values of μ and σ . The values of μ and σ completely determine the normal distribution. When μ and σ are known, the proportion of population values in any interval can be evaluated.
3. If σ remains fixed but μ changes, the density of the random variable remains the same *shape*, but its location changes.
4. If μ remains fixed but σ changes, the density of the normal random variable has the same *location* but its shape changes.

The 68-95-99.7 Rule

In the normal distribution with mean μ and standard deviation σ

- 68% of the observations fall within σ of the mean μ
- 95% of the observations fall within 2σ of the mean μ
- 99.7% of the observations fall within 3σ of the mean μ

Example of the 68-95-99.7 rule

The distribution of systolic blood pressure in 18- to 74-year-old males in the US is approximately normal with mean $\mu = 129$ mm of mercury and standard deviation $\sigma = 20$ mm of mercury.

The standard normal distribution

- The *standard normal distribution* is the normal distribution with
 - $\mu = 0$
 - $\sigma = 1$
- The name Z is often used for a variable that has the standard normal distribution.

For this particular normal distribution,

Using tables of the standard normal distribution

- What if we wanted to know what proportion of values of a standard normal variable Z were less than some particular value?
- Suppose we live in a particular Scandinavian city, where temperature is measured in Centigrade. Weather records kept for many years indicate that the temperature at 11:00 a.m. on Jan. 28 follows a standard normal distribution.

- We want to know in what proportion of years we can expect the temperature at this time to be less than or equal to -1.5 C.
- We could use Table A in your textbook.
 - The proportion is 0.0668.
- Similarly, the proportion of years we can expect the temperature at this time to be ≤ -1.75 C is .0401.

What if we instead wanted to know the proportion of years with temperature ≥ -1.75 ? (Remember that the total area under the normal curve is 1.0.)

What if we instead wanted to know the proportion of years with temperature $\geq +1.75$? (Remember symmetry of the normal distribution.)

Standardizing values from other normal distributions

All normal distributions would be the same if we measured in units of size σ around the mean μ as center!

If x is an observation from a distribution that has mean μ and standard deviation σ , the *standardized value* of x is

$$z = \frac{x - \mu}{\sigma}$$

Standardized values are often called *z-scores*.

z-scores tell how many standard deviations the original observation is away from the mean of the distribution, and in which direction.

- If the z-score is positive, the original observation was larger than the mean μ .
- If the z-score is negative, the original observation was smaller than μ .

Example of z-scores

Recall that the distribution of systolic blood pressure of men aged 18-74 is approximately normal with $\mu = 129$ mm Hg and $\sigma = 20$ mm Hg. The standardized height is

$$z = \frac{sbp - 129}{20}$$

If a man has sbp = 157 mm Hg, his standardized sbp is

$$z = \frac{157 - 129}{20} = 1.4$$

If a man has sbp = 93 mm Hg, his standardized sbp is

$$z = \frac{93 - 129}{20} = -1.8$$

General procedure for finding normal proportions

1. State the problem in terms of the observed variable X .
2. Standardize the value of interest x to restate the problem in terms of a standard normal variable Z . You may then wish to draw a picture to show the area under the standard normal curve.
3. Find the required area under the standard normal curve, using Table A and remembering
 - The total area under the curve is 1.0.
 - The normal distribution is symmetric.

Using the standard normal distribution to compute proportions for other normal distributions

Let's use the symbol X for a variable representing the systolic blood pressure of men. What proportion of men have sbp < 100?

If a man has sbp = 100, his standardized sbp is

$$z = \frac{100 - 129}{20} = -1.45$$

According to Table A, the proportion of values of a standard normal variable that are less than or equal to -1.45 is 0.0735.

This proportion is the same as the proportion of X values that will be less than 100.