List of Definitions, Concepts, and Formulas for Math 125

Definitions and Concepts

Chapter 3

Endpoint Convention for Continuous Data: For a particular number where the data are rounded, the interval generally goes from that number $- 1/2$ to that value $+ 1/2$. For example: 166 pounds (rounded) runs from 165.5 to 166.5 pounds.

For a value where the data are truncated, the interval goes from the number to the number $+1$. For example: 22 years (truncated assumed) runs from 22.000 to 22.999 years.

Endpoint Convention for Discrete Counting Data: Here each number has to have an interval one wide centered at that value. For that reason, the base is always $\pm 1/2$ from the number itself. For example: 3 children in family size daa will have an interval running from 2.5 to 3.5 children (see page 43 and the histogram on page 44).

Rule to Find the Area of a Block in a Hisogram: To find the area (or percent of the data) for a block, multiply the length of the interval by the density.

Types of Variables; Qualitative (descriptive) or quantitative (numeric); continuous (takes on all values in the interval) or discrete (takes on only certain, separated values).

Density Scale: The vertical scale for a histogram. It gives the percent of the area of the histogram per unit of the horizontal scale. It is a density scale, meaning the taller the block of the histogram, the more densely packed it is.

(Unfortunately, the educational example on pages 38 to 40 is very misleading. Please skip those pages completely. Example 1 on page 40 is OK.)

Chapter 4

Root-mean-square:

Do these steps in this order. The name of the operation is read backwards.

- SQUARE all the entries, getting rid of the signs.
- Take the MEAN (average) of the squares.
- Take the square ROOT of the mean.

The formula is

$$\text{r.m.s. size of a list} = \sqrt{\text{average of } \left(\text{entries}^2\right)}.$$

Chapter 5

Standard Units:

A value is converted to standard units by seeing how many SDs it is above or below the average.

> *For example: 2.3 SDs above average says that the standard units are +2.3;*
> *0.84 SDs below average means that the standard units are −0.84.*
>
> *Also, when the standard units are 1.36, the value is 1.36 SDs above average;*
> *when the standard units are −2.84, the value is 2.84 SDs below average.*
>
> *Finally, standard units of 0 means that the value is at the average.*

Finding Areas under the Normal Curve:

As an example of using the Normal Table of the text, consider finding the area between 1 and 2.

First find the half-areas between 0 and 1 and between 0 and 2. Each requires looking up the number and dividing the areas by 2: 68.27%/2 and 95.45%/2.

Then subtract the two half areas to find the area of the sliver: $47.725\% - 34.135\% = 13.59\%$.

(There is not a unique way to do this problem. Compare example 7 on page 84.)

Example 6 on page 84 asks to find the area to the left of 2.

The area to the left of 2 is the sum of the area to the left of 0, and the area between 0 and 2.

The area to the left of 0 is half the total area, by symmetry: $\frac{1}{2} \times 100\% = 50\%$.

The area between 0 and 2 is half the area between −2 and 2: 95.45%/2=47.725%.

The sum is $50\% + 47.725\% = 97.725\%$.

Example 8 on pages 85 and 86 is useful additional practice.

The standard units for 63 inches and 72 inches are immediately known to be −2 and 1, as the heights are exactly 2 SDs below the average and 1 SD above the average.

Find the half-areas for 2 and 1. From the previous problem, they are 95.45%/2 and 68.27%/2.

Finally, add the two half-areas to get the total area of this region. It is very helpful to draw the picture of the normal curve and make two horizontal scales, one each for the heights and for the standard units. Answer: $47.725\% + 34.135\% = 81.86\%$.

Another noteworthy example is example 9 on page 87. To find the standard units, the text uses the previously-unstated formula

$$(59 - 63.5)/3 = -1.5.$$

The author has an aversion to formulas, but this formula is always useful.

The area above −1.5 is the area from −1.5 to 0 added to the area above 0. Look up the half-area for −1.5 to 1.5: 86.64%/2 and add 50%. Answer: $43.32\% + 50\% = 93.32\%$.

Pecentiles: A percentile is a number in the distribution that has that percent of the distribution blow it. The percenile is not a percent, but a member of the distribution. The associated percent is called the percentile rank.

Percentiles and the Normal Curve: To find the percentile rank for a given member of he distribution, just convert it to standard units and determine the area under the normal curve below that point.

To find the standard units for a given percentile rank when the rank is greater than 50%, double the amount over 50% and look it up as an area in the normal table. For example: the rank of 91% gives $91\% - 50\% = 41\%$. Double 41% and look up the z for 82%. The answer is 1.35.

When the percentile rank is less than 50%, one way to approach it is to double that percentage and subtract from 100%. Look up the z for that area and attach a minus sign to it. For example: the rank of 21% gives 42%, then $100\% - 42\% = 58\%$. Look up 58% and attach a minus sign to it. The answer is -0.80.

Formulas

Drawing a Histogram

First convert the distribution of the population into percentages for each category. Use the formula:

percentage for a category = 100% × (number in the given category) / (number in the total population).

In a histogram the areas of the blocks represent percentages.

To figure out the height of a block over a class interval, divide the percent by the length of the interval.

In a histogram, the height of a block represents crowding—percentage per horizontal unit.

With the density scale on the vertical axis, the areas of the blocks come out in percent. The area under the histogram over an interval equals the percentage of cases in that interval. The total area is 100%.

The Average and the Median: measurements of the center.

- The average of a list of numbers equals their sum, divided by how many there are.
- The median of a histogram is the value with half the area to the left and half the area to the right. For a list of numbers, the median is the middle number or the average of the two middle numbers if there are an even number of numbers on the list.
- Median versus average:

  They are equal if the histogram is symmetric, but the average is larger if the histogram has a long right-hand tail, and the average is smaller if the histogram has a long left-hand tail.

Calculation of the standard deviation.

- Find the average.
- Find the deviations from average by subtracting the average from each entry.
- Find the root-mean-square of the deviations from average (by taking their squares, the average of these squares, and the square root of that average—be sure to do these steps in this exact order).

    The Root-mean-square: Square all entries, take the mean (average) of the squares, and take the square root of the mean.

      The term mean-square refers to a square of which we have taken the mean. The term root-mean-square refers to a mean-square of which we have taken the root. That means that the operations are done in reverse order of their positions in the word.

As a formula: r.m.s. size of a list $= \sqrt{\text{average of (entries}^2\text{)}}$.

Conversion to Standard Units

A value is converted to standard units by seeing how many SDs it is above or below the average.

The formula is: $\quad \text{standard units} = \dfrac{\text{observation}-\text{average}}{\text{SD}}.$

Finding Areas under the Normal Curve

An area from minus a value to plus the same value is read off from the Normal Table; other areas are found by making a sketch and expressing the desired area in terms of areas that may be found by using the Table.

The Normal Approximation for Data

If a histogram follows the normal curve, approximate areas may be found by converting the endpoints to standard units and finding the appropriate areas under the normal curve by using the Table.

Percentiles and the Normal Curve

A. For a percentile rank above 50%, subtract 50% from the rank and double the result. That will give you the area between $-z$ and z. Use the normal table to detemine the z's. Here the positive answer is the relevant one.

B. Given a positive z, to find the associated percentile rank, take half of the area between $-z$ and z and add to it the area below $z = 0$ (50%).

C. For a percentile rank below 50%, to find the associated z, double the rank and subtract from 100%. From the table get $-z$ to z for that answer. Choose the negative one here.

D. Given a negative z, to find the associated percentile rank, look up $-z$ to z, subtract that middle area from 100%, and then divide by 2 to get the tail area. The tail area of the left-hand tail gives the percentile rank.

Examples: (a) Percentile rank is 91%. $(91\% - 50\%) \times 2 = 82\%$ z is 1.35.

(b) $z = 0.20$. Half of area is $15.85\%/2 \approx 8\%$. $8\% + 50\% = 58\%$.

(c) Percentile rank is 5%. $100\% - (2 \times 5\%) = 100\% - 10\% = 90\%$. z is -1.65.

(d) $z = -0.50$. -0.50 to 0.50 is about 38% on the table. Subtract $100\% - 38\% = 62\%$. The tail area of the left-hand tail is $62\%/2 = 31\%$. That is the percentile rank.

Change of Scale

- Adding the same number to every entry on a list adds that constant to the average; the SD and the standard units do not change.
- Multiplying every entry on a list by the same positive number multiplies the average and the SD by that constant, but does not change the standard units.
- Multiplying all the numbers on a list by the same negative constant multiplies the average by that constant, multiplies the SD by the absolute value of that constant, and reverses the signs of the standard units.