7.4 from an x value to a probability p

Areas to the left of x

The probability p is the same as the area under the normal curve. Probability, expressed often as a percentage, is area. Probability is also the relative frequency. In this class probability, p, area, and relative frequency are all used interchangeably.

If x is not an whole number of standard deviations from the mean, then we cannot use a diagram as seen above. Spreadsheets have a function that calculates the area (probability) to the left of ANY x value. The letter p for probability is used for the area to the left of x.

The function that calculates the area to the left of x is:

=normdist(x,μ,σ,1)

The mean height μ for 43 female students in statistics is 62.0 inches with a standard deviation of 1.9. Determine the probability that a student is less than 60 inches tall (five feet tall).

In OpenOffice the probability p =
=normdist(60,62,1.9,1)
=0.1463

14.63% of the area is to the left of 60 inches. The probability a female student in statistics class is below 60 inches is 14.63%.

Notation note: In probability notation the above would be written p(x < 60) = 0.1463

When the words "less than, smaller, shorter, fewer, up to and including" are used then the NORMDIST function can be used to calculate the probability.

Area to the right of x
The mean number of cups of sakau consumed in sakau markets on Pohnpei is $\mu = 3.65$ with a standard deviation of $\sigma = 2.52$. Note that this data is actually based on customer data for 227 customers at four markets - one near Kolonia and three in Kitti. Although this data is actually sample data and not population data, we will treat the mean and standard deviation as population parameters. The data is not perfectly normally distributed. The data is, however, distributed in a reasonably smooth heap.

What is the probability a customer will drink more than five cups?

Note the word "more." If the question were "What is the probability that a customer will drink less than five cups, then the solution would be $=\text{NORMDIST}(5,3.65,2.52,1)$. This result is 0.70 or a 70% probability a customer will drink less than five cups.

The area under the whole normal curve is 1.00. Remember that 1.00 is also 100%. If 70% drink less than five cups, then we can calculate the probability that those who drink more than five cups is 30%. $100\% - 70\% = 30\%$.

Or $1.00 - 0.70 = 0.30$

Making a sketch of the normal curve including the mean, the $x$-value, and the area of interest can help determine when to subtract a result from one and when to not.

**Area between two $x$ values**

A study of the prevalence of diabetes in a village on Pohnpei found a mean fasting blood sugar level of $\mu = 117$ with a standard deviation $\sigma = 33$ in mg/dl for females aged 20 to 29 years old. Blood sugar levels between 120 and 130 are considered borderline diabetes cases. What percentage of the females aged 20 to 29 years old in this village are between a mean fasting blood sugar of 120 and 130 mg/dl?

For this example, presume that the distribution is normal.

The probability is the percentage. The probability is the area between $x = 120$ and $x = 130$ as seen in the image below.
In probability notation this would be written $p(120 < x < 130) =$ ?

To obtain the area between 120 and 130, calculate the area to the left of 120.

Then calculate the area to the left of 130.
Subtract the area to the left of 120 from the area to the left of 130. What remains is the area between 120 and 130.

The table below represents a spreadsheet laid out to calculate the area to the left of 120 in column B and the area to the left of 130 in column C.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>x</td>
<td>120</td>
<td>130</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>mean µ</td>
<td>117</td>
<td>117</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>stdev σ</td>
<td>33</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>normdist</td>
<td>=NORMDIST(B1,B2,B3,1)</td>
<td>=NORMDIST(C1,C2,C3,1)</td>
<td>=C4-B4</td>
</tr>
<tr>
<td>4</td>
<td>normdist</td>
<td>0.54</td>
<td>0.65</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Row four is presented twice: once with the formulas and once with the results of the formulas.

The area to the left of 120 is 0.54. The area to the left of 130 is 0.65. 0.65 − 0.54 is 0.11. The probability that females aged 20 to 29 years old in this village have a blood sugar level between 120 and 130 is 11%.