PERCENTAGES UNDER THE NORMAL DISTRIBUTION

AMS102

Notations. $N(\mu, \sigma)$ is the normal distribution with the mean μ and standard deviation σ.

The percentage of data lying between values a and b is denoted $P(a < x < b)$ ("P" stands for percentage or proportion).

The percentage of data lying above a is denoted $P(x > a)$; below b, $P(x < b)$.

z-score. Under the normal $N(\mu, \sigma)$, the z-score of the value x is $\boxed{z = \frac{x-\mu}{\sigma}}$. The process of computing the z-score is called standartization.

EXAMPLE. Consider the distribution $N(100, 10)$. The z-score of 100 is $\frac{100-100}{10} = 0$. The z-score of 105.3 is $\frac{105.3-100}{10} = 0.53$.

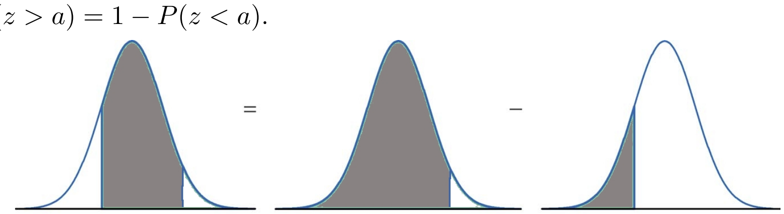
Finding the percentage. In order to compute percentages under a normal distribution, you need to standartize every given value. For example, to find $P(x < b)$ under the normal distribution $N(\mu, \sigma)$, you first standartize b to $\frac{b-\mu}{\sigma}$. Then you need to find $P(z < \frac{b-\mu}{\sigma})$. Look up the value of $\frac{b-\mu}{\sigma}$ in table A ("Standard normal probabilities"). The corresponding number in the table is the required proportion. To convert to percentages, multiply by 100%.

EXAMPLE, CONTINUED. Consider the normal distribution $N(100, 10)$. To find the percentage of data below 105.3, that is $P(x < 105.3)$, standartize first:

$$P(x < 105.3) = P\left(z < \frac{105.3 - 100}{10}\right) = P(z < 0.53).$$

Then find the proportion corresponding to 0.53 in Table A: look for the intersection of the row labeled 0.5 and the column labeled .03. The number is .7019. Thus $P(x < 105.3) = .7019$ or 70.19%.

Table A gives only proportions of the kind $P(z < b)$. To find other proportions, we use geometric facts that $P(a < z < b) = P(z < b) - P(z < a)$ (see the picture) and $P(z > a) = 1 - P(z < a)$.



EXAMPLE, CONTINUED. Consider the normal distribution $N(100, 10)$. To find $P(97.1 < x < 105.3)$, standartize first:

$$P(97.1 < x < 105.3) = P\left(\frac{97.1 - 100}{10} < z < \frac{105.3 - 100}{10}\right) = P(-0.29 < z < 0.53).$$

Then

$$P(-0.29 < z < 0.53) = P(z < 0.53) - P(z < -0.29).$$

The last two proportions can be found in Table A: $P(z < 0.53) = .7019$ and $P(z < -0.29) = .3859$ (row -0.2, column 0.09). Thus

$$P(97.1 < x < 105.3) = .7019 - .3859 = .3160 \text{ or } 31.6\%.$$

From percentages to values. There is another kind of problems: given a percentage, find the corresponding boundary value. For example, given the percentage $P(x < b) = P$, what is b? Here to find b, we look up P or the value closest to P in the table and find the corresponding z-score. Then, we need to solve $z = \frac{b-\mu}{\sigma}$ for b. Algebra shows that $b = z\sigma + \mu$.

Example, continued. Consider the normal distribution $N(100, 10)$. What values lie in the lower 80% of the data?

We need to find b such $P(x < b) = 80\%$. First we find the z-score Z such that $P(z < Z) = 80\%$. The table does not contain 0.8; the closest number is 0.7995. It lies in the row 0.8 and column 0.04. Thus the z-score of b is approximately 0.84:

$$0.84 = \frac{b - 100}{10}.$$

Hence $b - 100 = 0.84 \times 10 = 8.4$ and $b = 100 + 8.4 = 108.4$. We conclude that the lower 80% of this distribution is formed by values below 108.4