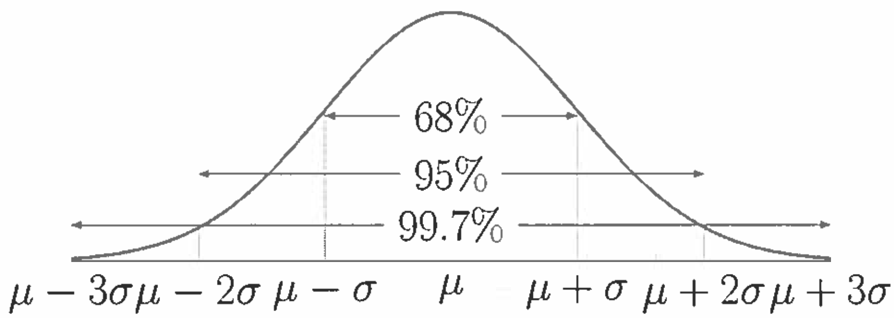
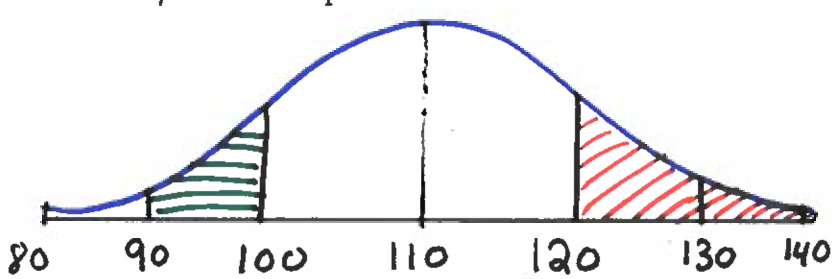
68-95-99.7 Rule for Normal Models (Empirical Rule)

In a Normal model, approximately:

- 68% of the values fall within one standard deviation of the mean.
- 95% of the values fall within two standard deviations of the mean.
- 99.7% of the values fall within three standard deviations of the mean.



Example: Suppose the U of A student IQs are (approximately) normally distributed with mean $\mu = 110$ points and standard deviation of $\sigma = 10$ points.

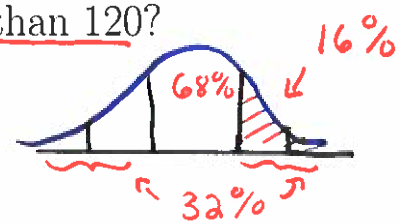


a) In what interval would you expect the central 95% of IQs to be found?

[90, 130]

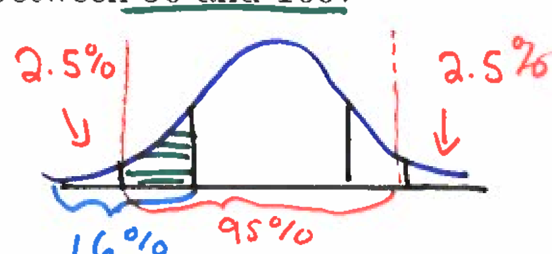
b) What percent of students will have an IQ of more than 120?

16%



c) What percent of students will have an IQ of between 90 and 100?

16% - 2.5% = 13.5%

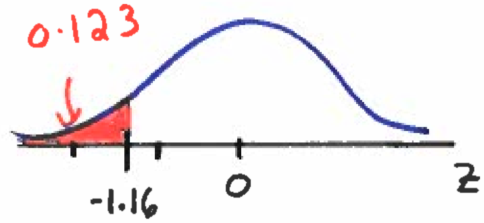


Section 5.4: Finding Normal Percentiles

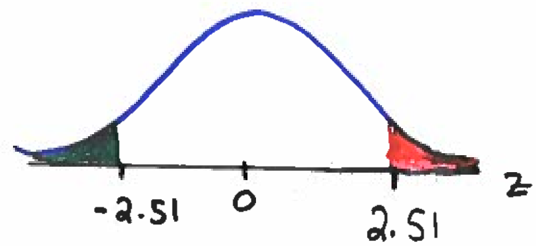
Standard Normal Distribution

The z-distribution table (Table Z) can be found in Appendix C of the textbook or on eClass.

Using the z- distribution, find



a) $P(z < -1.16)$ $= 0.1230$



b) $P(z > 2.51)$ or

$= P(z < -2.51)$

$= 0.006$

$P(z > 2.51)$

$= 1 - P(z < 2.51)$

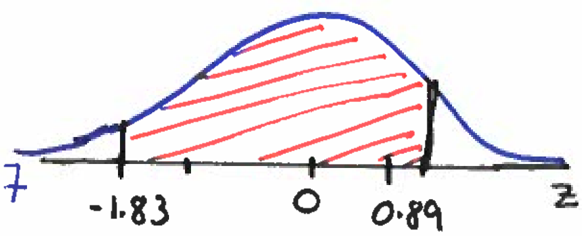
$= 1 - 0.9940$

$= 0.006$

c) $P(-1.83 < z < 0.89)$

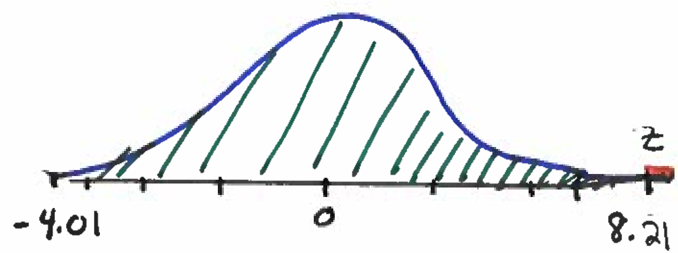
$= P(z < 0.89) - P(z < -1.83)$

$= 0.8133 - 0.0336 = 0.7797$



d) $P(z > 8.21)$ $\simeq 0$

e) $P(z > -4.01)$ $\simeq 1$

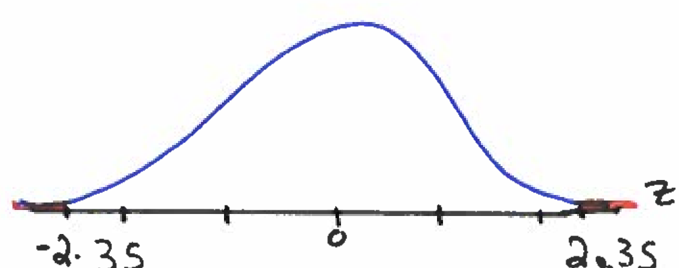


f) the probability that an outcome falls more than 2.35 standard deviations above or below the mean.

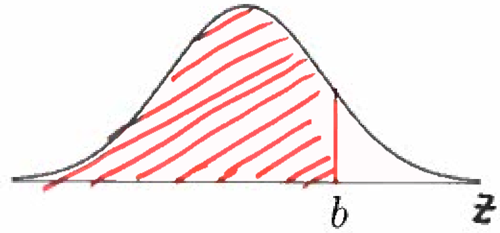
$P(z > 2.35 \text{ or } z < -2.35)$

$= 2P(z < -2.35)$

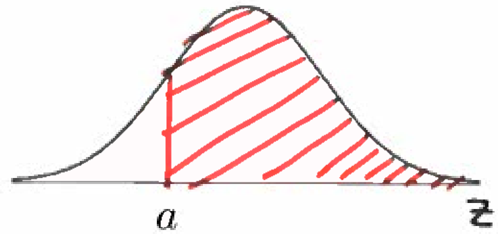
$= 2(0.0094) = 0.0188$



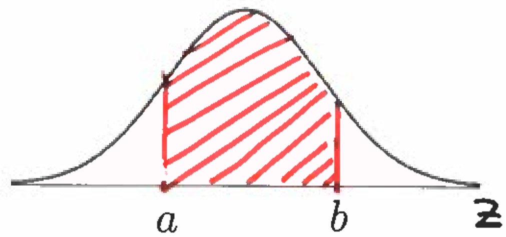
$P(z < b)$



$P(z > a)$
$= 1 - P(z < a)$
$= P(z < -a)$



$P(a < z < b)$
$= P(z < b) - P(z < a)$



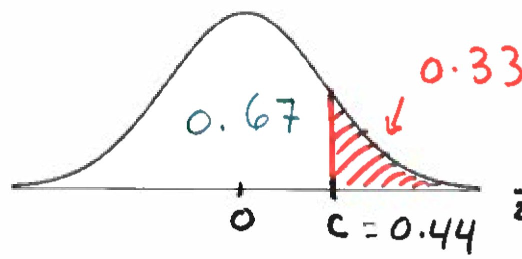
In some cases, we will be given the area to be captured under the z-curve. We would then have to find the interval that captures this area.

Example: Find the value of c for which

a) $P(z > c) = 0.33$

$\therefore$ $c = 0.44$

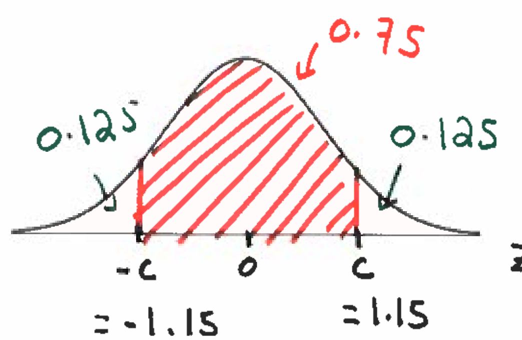
$(P(z > 0.44) = 0.33)$



b) $P(-c < z < c) = 0.75$

$\therefore$ $c = 1.15$

$(P(-1.15 < z < 1.15) = 0.75)$



Note: If the exact value of the area you are looking for is not in the z table, then choose the closest entry in the table. or if value is half way between two entries, take the average of the two z values.

Other Normal Distributions

percentiles

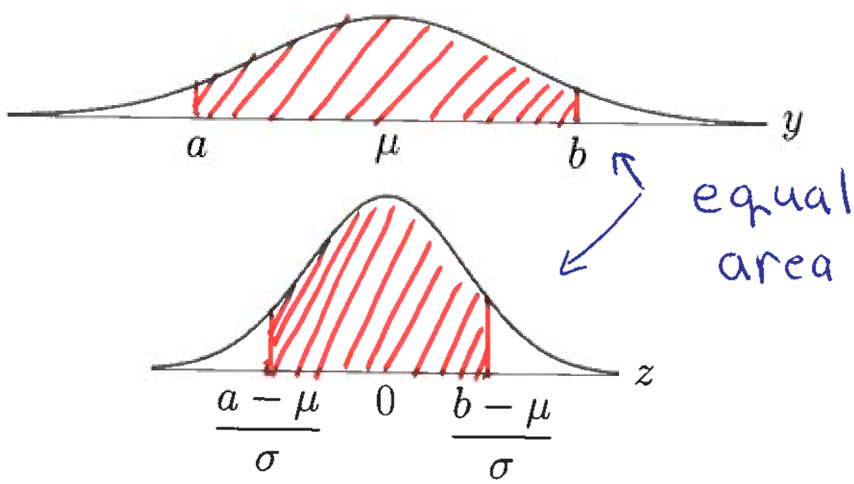
To calculate probabilities/proportions for other Normal distributions, we standardize to the standard Normal distribution and use the z curve to do the calculations.

Key Fact:

If y is Normally distributed with mean μ and standard deviation σ (so that y can be described by the Normal model $N(\mu, \sigma)$), then the standardized variable

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

has the standard Normal distribution (so that z can be described by the Normal model $N(0, 1)$). This process of standardizing is area-preserving:



Thus,

- $P(y < b) = P\left(z < \dfrac{b - \mu}{\sigma}\right)$
- $P(y > a) = P\left(z > \dfrac{a - \mu}{\sigma}\right)$
- $P(a < y < b) = P\left(\dfrac{a - \mu}{\sigma} < z < \dfrac{b - \mu}{\sigma}\right)$

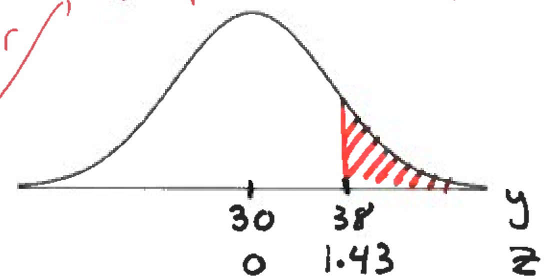
To convert z-scores to y-values, we can use the formula:

$$y = \sigma z + \mu$$

Example: The lifetime of a particular brand of flashlight battery is (approximately) normally distributed with a mean of 30 hours and a standard deviation of 5.6 hours. $y =$ battery life $\sim N(30, 5.6)$

a) What is the probability that a randomly selected battery will last longer than 38 hours?

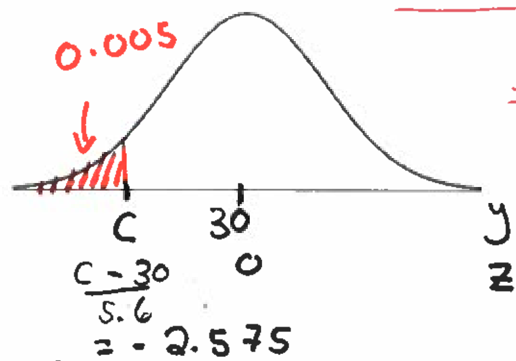
$$P(y > 38) = P\left(z > \frac{38-30}{5.6}\right) \text{ or } 1 - P(z < 1.43)$$
$$= P(z > 1.43)$$
$$= P(z < -1.43)$$
$$= 0.0764$$



b) Suppose the manufacturer will give a refund to any customer who has purchased a battery with a lifespan in the bottom 0.5%. What is the maximum battery life that qualifies for a refund?

$$\frac{-2.57 + (-2.58)}{2} = -2.575$$

$$P(y < c) = 0.005$$
$$\Leftrightarrow P\left(z < \frac{c-30}{5.6}\right) = 0.005$$
$$\Leftrightarrow \frac{c-30}{5.6} = -2.575$$
$$\Leftrightarrow c = 5.6(-2.575) + 30$$
$$\approx 15.58 \text{ hours}$$

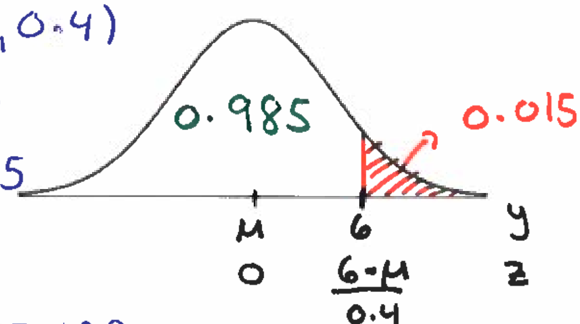


Example: A paint tinting machine can be set so that it dispenses an average of μ ml of dye per can of paint. Suppose that the amount of dye dispensed is (approximately) normally distributed with a standard deviation of 0.4 ml. If more than 6 ml of dye are dispensed when mixing a certain colour, then the shade is not acceptable. Find the setting for μ so that only 1.5% of the cans of paint are not acceptable.

$y =$ dye dispense amount $\sim N(\mu, 0.4)$

Find μ so that $P(y > 6) = 0.015$

$$\Leftrightarrow P\left(z > \frac{6-\mu}{0.4}\right) = 0.015$$
$$\Leftrightarrow P\left(z < \frac{6-\mu}{0.4}\right) = 0.985$$
$$\Leftrightarrow \frac{6-\mu}{0.4} = 2.17 \Leftrightarrow \mu = 5.132 \text{ ml}$$



Example: Suppose that the scores on a national university entrance exam are (approximately) normally distributed with a mean of 500 and a standard deviation of 100. $y =$ exam score $\sim N(500, 100)$

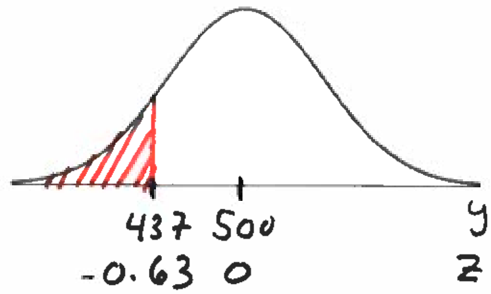
a) What is the probability that a randomly selected student scored less than 437?

$P(y < 437)$

$= P\left(z < \frac{437-500}{100}\right)$

$= P(z < -0.63)$

$= 0.2643$



b) What is the probability that a randomly selected student scored between 280 and 675?

$P(280 < y < 675)$

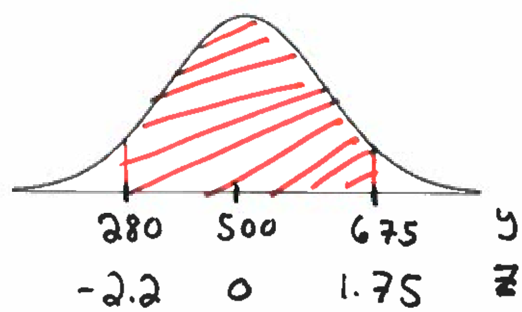
$= P\left(\frac{280-500}{100} < z < \frac{675-500}{100}\right)$

$= P(-2.2 < z < 1.75)$

$= P(z < 1.75) - P(z < -2.2)$

$= 0.9599 - 0.0139$

$= 0.946$



c) Suppose that a certain university only admits students who score within the top 14%. What is the minimum score on the exam that a student can have to be eligible for admission at this university?

$P(y > c) = 0.14$

$\Leftrightarrow P(y < c) = 0.86$

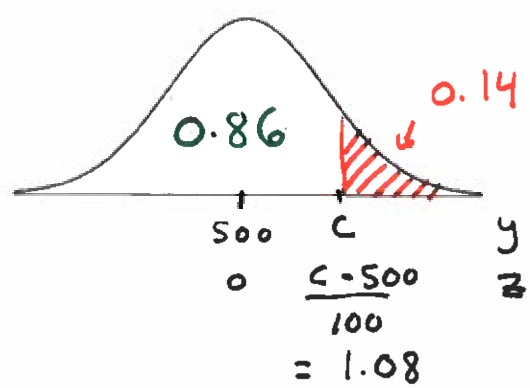
$\Leftrightarrow P\left(z < \frac{c-500}{100}\right) = 0.86$

$\Leftrightarrow \frac{c-500}{100} = 1.08$

$\Leftrightarrow c = 100(1.08) + 500$

$= 608$

$\therefore$ the min score is 608



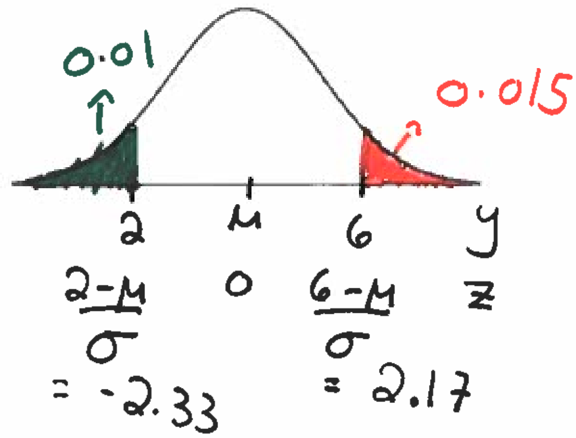
Example: Paint Machine Part II: $y = \text{dye dispense amount} \sim N(\mu, \sigma)$

Suppose a paint tinting machine can be set so that it dispenses an average of μ ml of dye per can of paint with a standard deviation of σ ml. Suppose that the amount of dye dispensed is (approximately) normally distributed. If more than 6 ml or less than 2 ml of dye are dispensed when mixing a certain colour, then the shade is not acceptable. Find the setting for μ and σ, so that only 1.5% of the cans of paint have more than 6 ml of dye and only 1% of the cans have less than 2 ml of dye.

Find μ and σ so that

$$P(y<2) = 0.01$$

and $P(y>6) = 0.015$



$P(y<2) = 0.01$

$\Leftrightarrow P\left(z < \frac{2-\mu}{\sigma}\right) = 0.01$

$\Leftrightarrow \frac{2-\mu}{\sigma} = -2.33$

$\Leftrightarrow -2.33\sigma + \mu = 2$

and $P(y>6) = 0.015$

$\Leftrightarrow P(y<6) = 0.985$

$\Leftrightarrow P\left(z < \frac{6-\mu}{\sigma}\right) = 0.985$

$\Leftrightarrow \frac{6-\mu}{\sigma} = 2.17$

$\Leftrightarrow 2.17\sigma + \mu = 6$

Solve:

$$\begin{array}{rl} & -2.33\sigma + \mu = 2 \\ - & 2.17\sigma + \mu = 6 \\ \hline & -4.5\sigma = -4 \end{array}$$

$\therefore \sigma = \frac{4}{4.5} \simeq 0.89$

$2.17\left(\frac{4}{4.5}\right) + \mu = 6$

$\therefore \mu = 6 - 2.17\left(\frac{4}{4.5}\right)$

$\simeq 4.07$

$\therefore \mu \simeq 4.07$ mL and $\sigma \simeq 0.89$ ml.

Chapter 11: From Randomness to Probability

A **random phenomenon** is any activity or situation in which there is uncertainty about which of two or more possible outcomes will result.

- know possible outcomes (values measured, observed, or reported)
- do not know which outcome will occur → Chance / statistical experiment

Example:

- Flipping a coin.
- "Hockey liking" status of a randomly selected person.
- Method of payment used by next customer.

A **trial** is a single attempt at a random phenomenon or a single occasion in which we observe a random phenomenon.

The **sample space** of a random phenomenon, denoted S, is the set of all possible outcomes of the random phenomenon. → Simple events

Example:

- Rolling a die once: (standard 6-sided) $S = \{1, 2, 3, 4, 5, 6\}$
- Flipping a coin once: $S = \{H, T\}$
- Flipping a coin twice: $S = \{HH, HT, TH, TT\}$

The sample space could be very large or it could be infinite.

- Rolling a die 10 times: 10 trials, 6 outcomes per trial $|S| = 6^{10}$ (# of elements in S)
  (2, 2, 6, 1, 5, 3, 4, 1, 1, 2)
- Flipping a coin until we get a tail:

$S = \{T, HT, HHT, HHHT, \ldots\}$

To illustrate a sample space, we can use → table

- **Venn diagrams**

  Example: Flipping a coin three times:

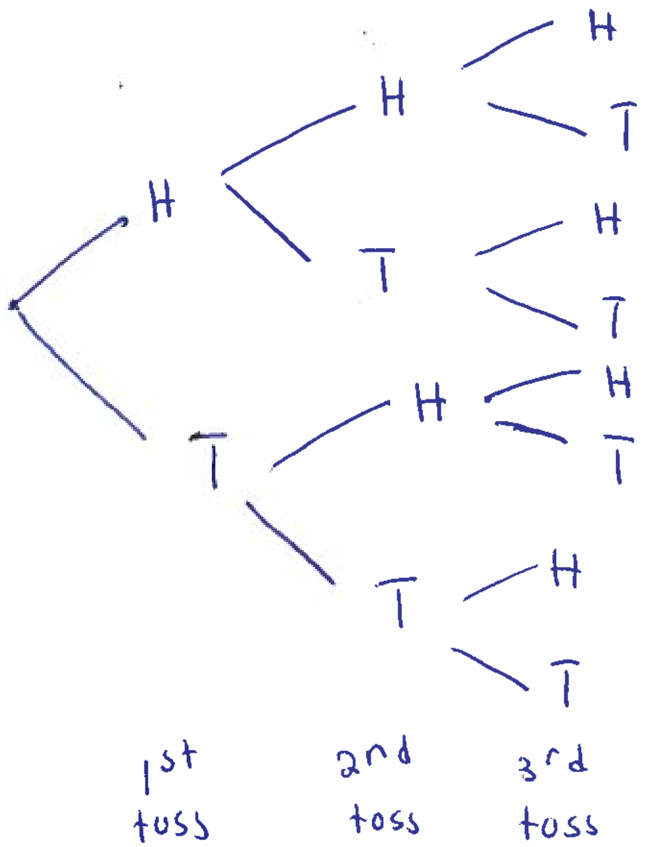
  HHT $\quad HHH \quad$ TTH

  HTH $\quad$ THT

  THH $\quad TTT \quad$ HTT

- **Tree diagrams**

  Example: Flipping a coin three times:



Outcomes

HHH

HHT

HTH

HTT

THH

THT

TTH

TTT