Normal model

Date: 2017-09-07

What it is

The normal model is the ubiquitous "bell-shaped curve" that is one example of a model for continuous data. The normal model is symmetric and peaks at the mean.

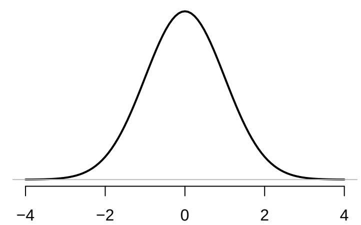

Figure 1: A normal model with mean 0 and standard deviation 1.

When to use it

The normal model is a common one in statistical inference. The CLT shows that for large samples, the totals and averages of observations follow approximately a normal model (see *Why it works*). In addition, many biological and psychological measurements (height, IQ, test scores) follow normal models.

How to use it

The **Empirical Rule** can be useful to make quick estimates. It states that for a normal model, 68% of the probability falls within 1 standard deviation of the mean, 95% falls within 2 standard deviations of the mean, and 99.7% falls within three standard deviations of the mean.

To compute normal probabilities in R, use the `pnorm()` function. If Y is modeled by a normal model with mean m and standard deviation s, then $P(Y \leq y) =$ `pnorm(q=y, mean=m, sd=s)`. Multiple `pnorms` can be used to find other types of areas. For example, say the mean is $m = 100$ and standard deviation is $s = 15$[1]. Then

[1] This is approximately the distribution of IQ scores in adults

```
# P(90 < Y < 120)
pnorm(q = 120, mean = 100, sd = 15) - pnorm(q = 90,
    mean = 100, sd = 15)
```

```
## [1] 0.6562962
```

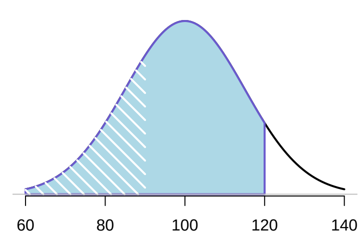

Figure 2: P(90 < Y < 120) = P(Y < 120) - P(Y < 90).

```
# P(Y > 130)
1 - pnorm(q = 130, mean = 100, sd = 15)
```

```
## [1] 0.02275013
```

Why it works