	Bayesians	Frequentists
probability is	subjective (but not arbitrary)	objective
it is	degree of belief	long-run frequency
it is in	in beliefs of rational agents	in the world
it applies to	anything you can believe in	only to repeatable events

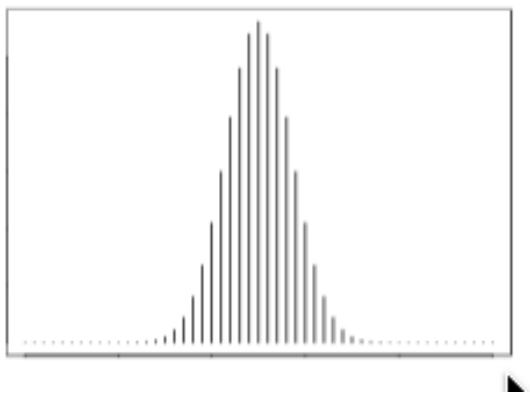
Probabilities go from 0 to 1
And must sum to 1!

Playing with distributions in R

	Definition	Binomial	Normal
d	Probability of specifc outcome	x=outcome, size,prob	NA
p	Chance outcome doesn't exceed threshold	q=threshold, size,prob	q=threshold, mean,sd
q	Some quantile	p=quantile, size,prob	p=quantile, mean,sd
r	Sample random numbers	r=#of draws, size,prob	r=# of draws, mean,sd

Two important distributions

Binomial



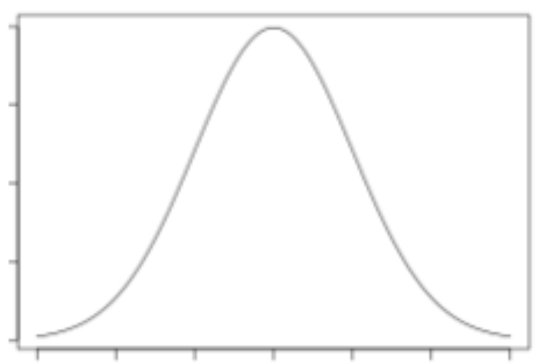
Count data with two outcomes (coin toss)

Discrete

size = # of trials

prob = "weight" of coin

Normal



"Bell curve"

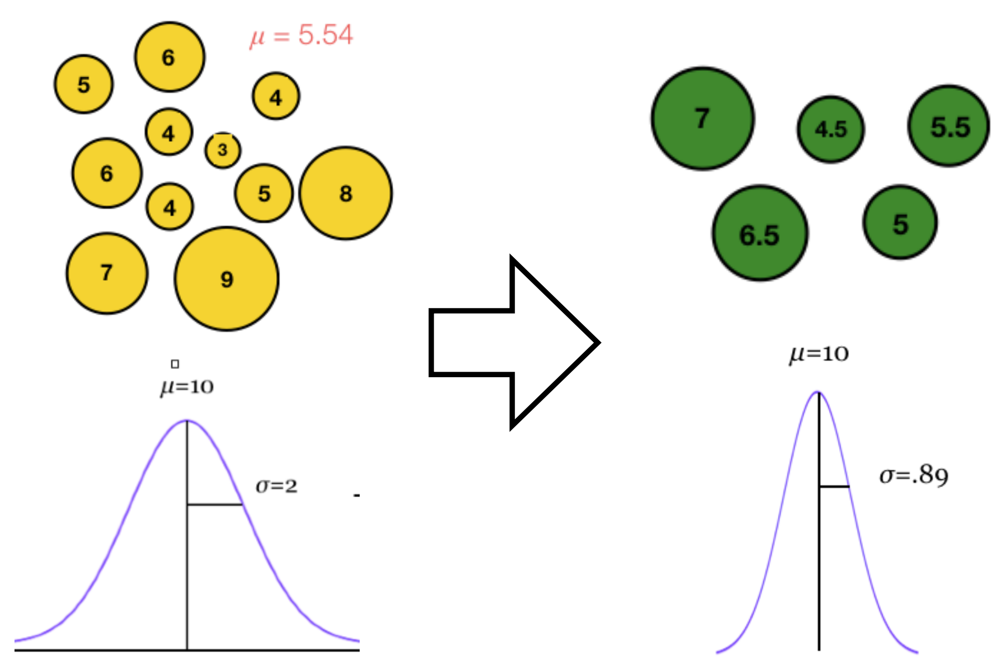
Continuous

mean = centre of mass

sd = measure of spread

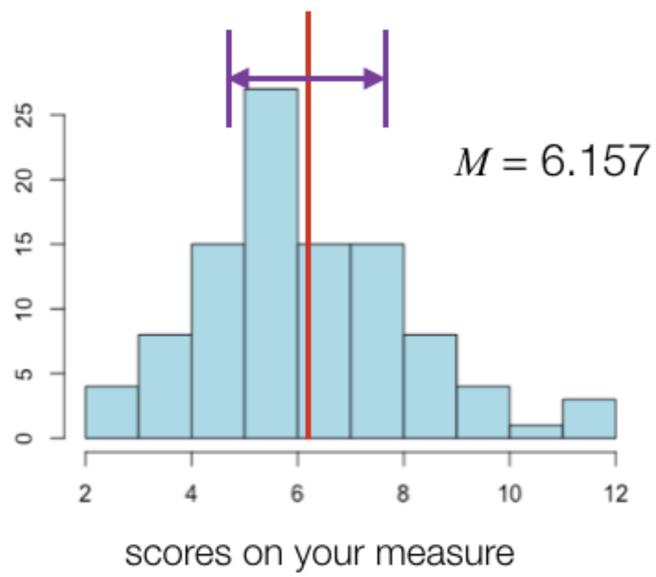
thing	"usual" symbol	thing	"usual" symbol	what is it?	do we know its value?
true population mean	μ	**true population sd**	σ	the truth	no
estimated population mean	$\hat{\mu}$	**estimated population sd**	$\hat{\sigma}$	a statistical inference	yes
sample mean	$\bar{X}$ or M	**sample sd**	s	a description of our dataset	yes

Sampling distribution of the mean is a theoretical idea that captures what you would expect the means of lots of samples from a population to look like



It is less variable than the original distribution

- **Central limit theorem:** As sample size goes to infinity, the sampling distribution of the mean becomes normal, regardless of the shape of the underlying distribution



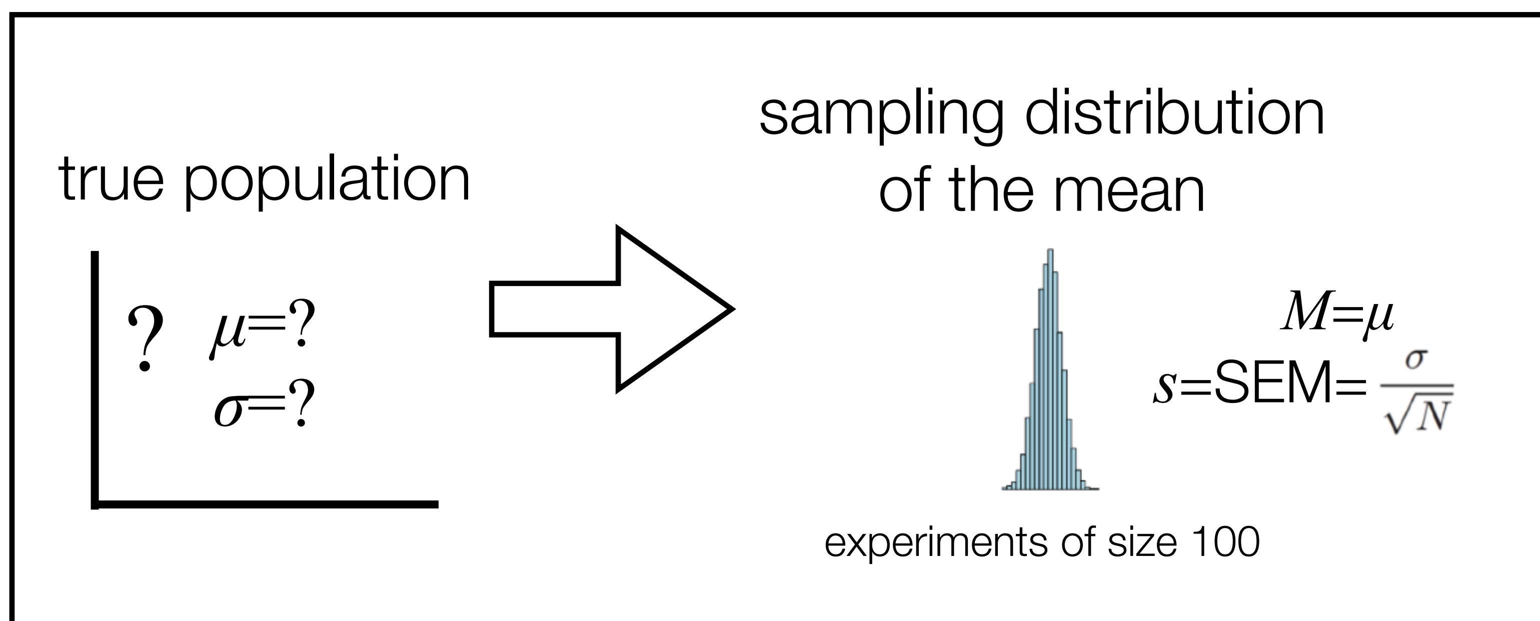
the 95% confidence interval (CI) is the range that covers the mean 95% of the time

$$\mathrm{CI}_{95} = \bar{X} \pm 1.96 \frac{\hat{\sigma}}{\sqrt{N}}$$

```
library(lsr)
ciMean(x)

ciMean( data )

ciMean( x,conf=.8 )
```



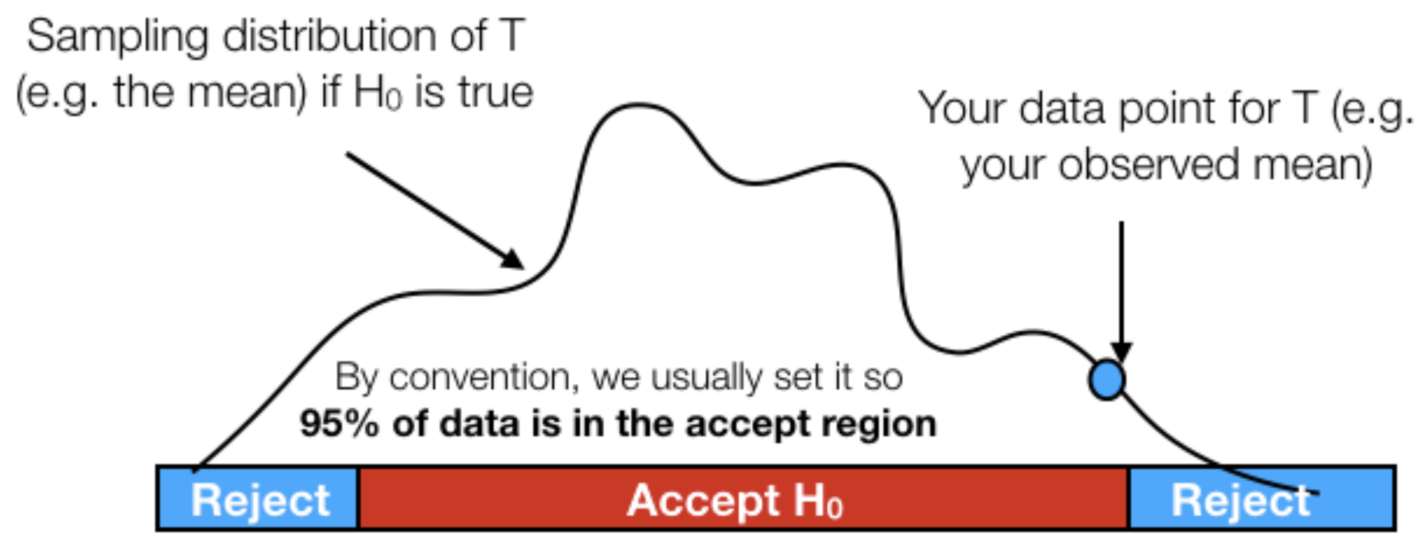
Research	Implicit other research hypothesis "people disagree with each other about whether we're running out of food" = YES / NO	Hypothesis scientist cares about "people agree with each other about whether we're running out of food" != YES / NO
Statistical	Exact opposite of the alternative statistical hypothesis $H_0 : \theta = 0.5$ (the null)	Statistical consequences of research hypothesis $H_1 : \theta \neq 0.5$ (the alternative)

	accept H_0	reject H_0
H_0 is true	Correct!	Type I error: false positive α Type I error rate significance level
H_0 is false	Type II error: false negative β Type II error rate	Correct! $1-\beta$ power

Things you can say	**Things you can't say**
p is the Type 1 error rate you are willing to tolerate to reject H0	p is the probability the null hypothesis is true
p is the probability, if H0 is true, of observing a test statistic at least as extreme as yours	the null hypothesis is true / false
we reject / retain the null / alternative hypothesis	p is the probability that the result was due to chance
the test was / was not significant	

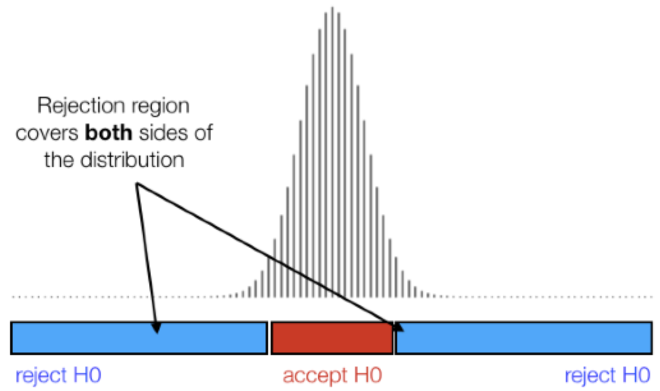
Designing a statistical test requires:

1) A diagnostic test statistic, *T* (e.g., mean)
2) Sampling distribution of *T* if the null is true
3) The observed *T* in your data
4) A rule that maps every value of *T* onto a decision (accept or reject H0)



Two kinds of tests vary by whether the rejection region is on both sides of the distribution (depends on your research hypothesis)

Two-tailed



One-tailed