Inference about a Mean

Sampling Distribution of Means:

You want to know the mean of a population. But you realize the population is too large to actually compute the true mean. So you observe a sample from the population and use the sample mean as an estimate of the population mean. You realize the sample mean would change with another sample, so you decide to make statistical inference about the population mean based on information in the sample.

The population from which you collect data is normally distributed with mean μ and standard deviation σ. Let y denote an observation from the population.

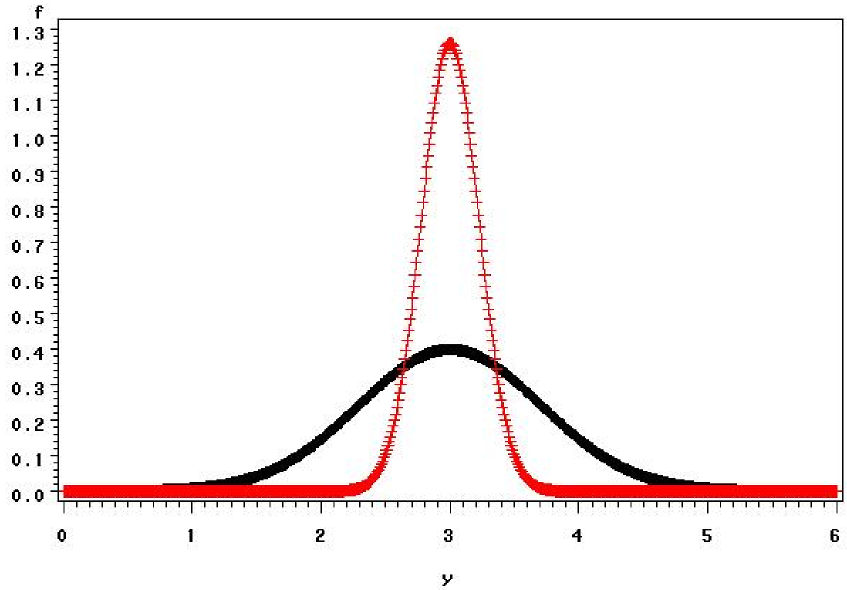
Draw sample of size n → $y_1, y_2, \cdots y_n$.

Compute the *sample* mean $\overline{y} = \left(y_1 + \cdots + y_n\right) / n$.

Then sampling distribution of $\overline{y}$ is normal with mean μ and standard deviation $\sigma_{\overline{y}} = \sigma / \sqrt{n}$.

Sampling Distribution of Means from a Normal Distribution

Population distribution of y and sampling distribution of $\bar{y}$



Black curve: Distribution of the population
Mean = μ
Standard deviation = σ

Red curve: Sampling distribution of $\bar{y}$
Mean = μ
Standard deviation = $\sigma / \sqrt{n}$

Empirical Rule applied to Sampling Distribution:

95% of the time $\bar{y}$ is within $2\sigma / \sqrt{n}$ of μ

Confidence Interval for a Mean

Equivalent statements from the Empirical Rule:

95% of the time μ is within $2\sigma / \sqrt{n}$ of $\overline{y}$

95% of the time μ is in the interval $(\overline{y} - 2\sigma / \sqrt{n}, \overline{y} + 2\sigma / \sqrt{n})$

This is a 95% Confidence Interval: $\left(\overline{y} - 2\sigma_{\overline{y}}, \overline{y} + 2\sigma_{\overline{y}}\right)$

Example:

Egg weight data: $\overline{y}$ = 65.4, s = 5.17, n=54

We don't know σ, so we use s in it's place. This is ok if n > 30.

$$s / \sqrt{n} = 5.17 / \sqrt{54} = .70 \qquad 2s / \sqrt{n} = 2(.70) = 1.40$$

$$\overline{y} \pm 2s / \sqrt{n} = 65.4 \pm 2(.70) = 65.4 \pm 1.40 \rightarrow (64.0, 66.8)$$

Interpretation of Confidence Interval

We are 95% confident that the population mean μ is in the interval (64.0, 66.8) in the following sense:

The population mean μ will be in the interval $\left(\bar{y} - 2\sigma_{\bar{y}}, \bar{y} + 2\sigma_{\bar{y}}\right)$ whenever $\bar{y}$ is within $2\sigma / \sqrt{n}$ of μ. From the Empirical Rule applied to the sampling distribution of $\bar{y}$, we know this happens 95% of the time.

Confidence Interval for a Mean

Conditions for interval $\bar{y} \pm 1.96\sigma / \sqrt{n}$ to be exactly valid:

1. Population normal, mean = μ, std. dev. = σ

   and

2. $y_1, y_2, \cdots y_n$ is a random sample from the population

The interval $\bar{y} \pm 2s / \sqrt{n}$ approximately valid if

1. Population approximately normal

and

2. n > 30

Test of Hypothesis about Mean

Suppose the long term mean for the egg weights is known to be μ=65. The mean of your current sample of 54 egg weights is 65.4.

Is there statistical evidence in the egg weight data that the population mean, μ, has changed? That is, is there statistical evidence that the sample mean 65.4 differs significantly from the hypothetical population mean of 65?

The answer to this question is "No," because the hypothetical mean 65 is contained in the 95% confidence interval:

$$65.4 - 2(.70) < 65 < 65.4 + 2(.70)$$

Equivalently: $|65.4 - 65|/.70 < 2$.

This is basically an example of a *Test of Hypothesis*. You are making a computation to check if the absolute difference between the *observed mean* and *hypothetical mean*, |65.4 – 65|, is greater than two standard errors of the mean, .70. The computation

$$|65.4 - 65|/.70$$

is the value of a *test statistic*.

Setup for Test of Hypothesis about a Mean

You want to test if there is evidence in your data that the population mean, μ, is different from a hypothesized value, μ_0.

Null Hypothesis: $H_0 : \mu = \mu_0$

Alternative Hypothesis: $H_0 : \mu \neq \mu_0$

Test Statistic: $z = (\bar{y} - \mu_0) / \sigma_{\bar{y}}$

Rejection Rule: Reject H_0 if $z < -2$ or $z > 2$ (equivalently, $|z| > 2$).

(The value 2 is a rounding of 1.96)

Margin of Error

Sampling distribution of $\overline{y}$ and its application:

95% of the time $\overline{y}$ will be within $2\sigma / \sqrt{n}$ of μ.

Stated another way:

95% of the time, the absolute difference between $\overline{y}$ and μ will be less than $2\sigma / \sqrt{n}$.

The probable “margin of error” of $\overline{y}$ as an estimate of μ is $2\sigma / \sqrt{n}$.

Margin of Error, "*MOE*"

Margin of Error: $MOE = 2\sigma / \sqrt{n} = 2\sigma_{\bar{y}}$

Relationship to confidence intervals:

95% of the time, μ will be between $\bar{y} - MOE$ and $\bar{y} + MOE$

Relationship to tests of hypothesis:

Reject the hypothesis that μ_0 is the true mean if $|\bar{y} - \mu_0| > MOE$.

Review: Inference about a population mean

Confidence interval: $\left(\bar{y} - 2\sigma_{\bar{y}}, \bar{y} + 2\sigma_{\bar{y}}\right)$

Example: 95% C. I. for mean egg weight

$\bar{y}$ = 65.4, s = 5.17,

n = 54 (use s in place of σ because n > 30)

MOE = $2s_{\bar{y}}$ = 1.40

$\bar{y} - 2s_{\bar{y}} = 63.99$ $\bar{y} + 2s_{\bar{y}} = 66.81$

$\Rightarrow$ 95% C. I. (64.0, 66.8)

Review: Inference about a population mean

Statistical test of hypothesis about μ :

$$H_0 : \mu = \mu_0 \qquad H_a : \mu \neq \mu_0$$

Test statistic: $$z = \frac{\bar{y} - \mu_0}{\sigma_{\bar{y}}} = \frac{\bar{y} - \mu_0}{\sigma / \sqrt{n}}$$

Reject H_0 if $|z| > 2$.

Example: Test whether mean is $\neq$ 65

$MOE = 1.40$

$| \bar{y} - \mu_0 | = .4 < MOE$

Equivalently,

$$z = \frac{\bar{y} - \mu_0}{\sigma / \sqrt{n}} = \frac{65.4 - 65}{.704} = \frac{.4}{.704} = .568 < 2$$

Conclude no statistical evidence that μ differs from $\mu_0 = 65$.

Type I and Type II errors of statistical test

Type I error : Reject true H_0

Type II error: Not reject false H_0

Analogy of statistical test to criminal trial

Test of Hypothesis

	State of Null Hypothesis	
Test Outcome	H_0 True	H_0 False
Reject	Type I error	OK
Not Reject	OK	Type II error

Criminal Trial

	State of Defendant	
Trial Outcome	Defendant Innocent	Defendant Guilty
Guilty	Convict innocent person	OK
Innocent	OK	Release guilty person

Diagnostic Test

	State of Subject	
Test Outcome	Uninfected (-)	Infected (+)
(+)	False positive	OK
(-)	OK	False negative

Error Probabilities in Hypothesis Testing

Probability of Type I error

α = .05 = P(Reject true H_0) = P(Type I error)

Probability of Type II error

β = P(Not reject false H_0) = P(Type II error) (depends on μ)

β (at $\mu = \mu_0$):

$$\beta = P\left(z < z_{\alpha/2} - \frac{|\,\mu_a - \mu_0\,|}{\sigma / \sqrt{n}} \right)$$

Computing Type II Error Probabilities

Ex. Type II error of egg weight test

$$H_0 : \mu = 65 \qquad H_a : \mu \neq 65$$

β (at μ =64) =

$$P\left(z < z_{\alpha/2} - \frac{|64-65|}{\sigma/\sqrt{n}} \right) = P\left(z < 1.96 - \frac{1}{.704} \right) = P\left(z < 1.96 - 1.42\right)$$

$= P(z < .54) = .7054$

β (at μ =63) =

$$P\left(z < z_{\alpha/2} - \frac{|63-65|}{\sigma/\sqrt{n}} \right) = P\left(z < 1.96 - \frac{2}{.704} \right) = P\left(z < 1.96 - 2.84\right)$$

$= P(\text{z} < -.88) = .1895$

β (at μ =62) =

$$P\left(z < z_{\alpha/2} - \frac{|62-65|}{\sigma/\sqrt{n}} \right) = P\left(z < 1.96 - \frac{3}{.704} \right) = P\left(z < 1.96 - 4.26\right)$$

$= P(\text{z} < -2.30) = .0108$

β (at μ =61) =

$$P\left(z < z_{\alpha/2} - \frac{|61-65|}{\sigma/\sqrt{n}} \right) = P\left(z < 1.96 - \frac{4}{.704} \right) = P\left(z < 1.96 - 5.68\right)$$

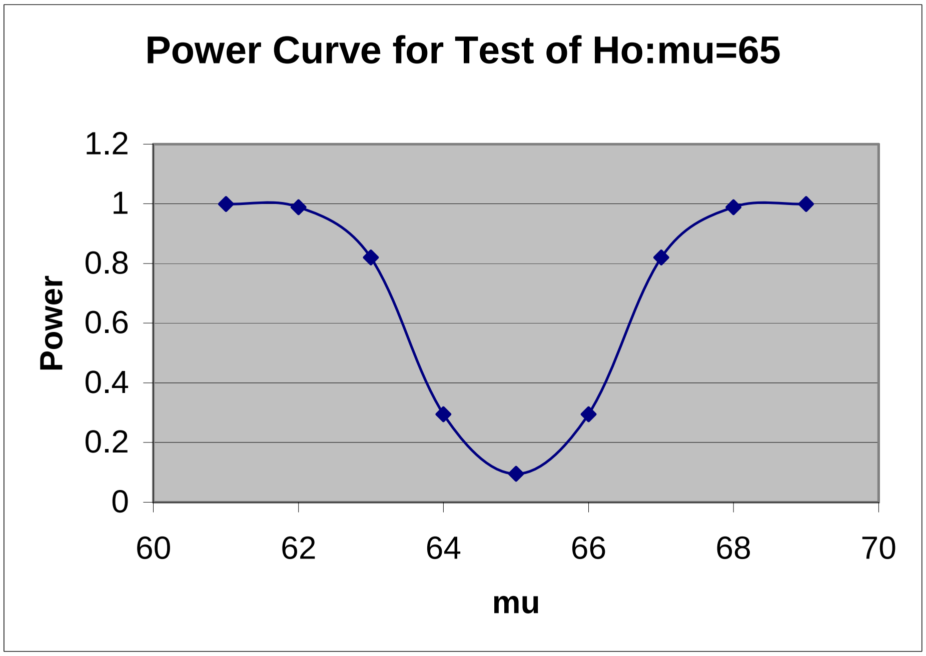
$= P(\text{z} < -3.72) = .0001$

Power of Statistical Test

Power = P(Reject false H_0) = 1- β

Example: Power of egg weight test

Power (at μ =61) = 1- β (at μ = 61) = 1-.0001 = .9999
Power (at μ =62) = 1- β (at μ = 62) = 1-.0108 = .9892
Power (at μ =63) = 1- β (at μ = 63) = 1-.1895 = .8105
Power (at μ =64) = 1- β (at μ = 64) = 1-.7054 = .2946
Power (at μ =65) = 1- β (at μ = 65) = .05
Power (at μ =66) = 1- β (at μ = 66) = 1-.7054 = .2946
Power (at μ =67) = 1- β (at μ = 67) = 1-.1895 = .8105
Power (at μ =68) = 1- β (at μ = 68) = 1-.0108 = .9892
Power (at μ =69) = 1- β (at μ = 69) = 1-.0001 = .9999



Sample Size

Sample size required to make MOE = E:

$$n = \frac{\sigma^2 \left(z_{\alpha/2}\right)^2}{E^2}$$

Sample size required to make Power of test equal to 1- β for $|\mu - \mu_0| = \Delta$:

$$n = \frac{\sigma^2 \left(z_{\alpha/2} + z_{\beta}\right)^2}{\Delta^2}$$

Sample Size Computation for MOE

Example: How many eggs would be needed in a sample to have MOE = .5?

$$n = (5.17)^2(1.96)^2/(.5)^2 = 4(26.7)^2/(.5)^2 = 427$$

Example: How many eggs to have MOE = 1?

$$n = (5.17)^2(1.96)^2/1 = 4(26.7)^2 = 107$$

Sample Size Computation for Power of Test

Example: How many eggs would be needed in a sample to obtain a test of hypothesis with power = .8 at μ = 64?

This means β = .2, Δ = 1, z_{β} = .81

$$n = (26.7)(1.96 + .81)^2 / 1 = (26.7)(2.77)^2 = 205$$

Other Topics in Hypothesis Testing

One-Sided Hypothesis Testing Problems

$$H_0 : \mu \le \mu_0 \quad \text{versus} \quad H_0 : \mu > \mu_0$$

Test statistic: $z = \dfrac{\overline{y} - \mu_0}{\sigma_{\overline{y}}} = \dfrac{\overline{y} - \mu_0}{\sigma / \sqrt{n}}$ Reject if $z > z_\alpha$.

$$H_0 : \mu \ge \mu_0 \quad \text{versus} \quad H_0 : \mu < \mu_0$$

Test statistic: $z = \dfrac{\overline{y} - \mu_0}{\sigma_{\overline{y}}} = \dfrac{\overline{y} - \mu_0}{\sigma / \sqrt{n}}$ Reject if $z < -z_\alpha$.

Other Topics in Hypothesis Testing

Power of one-sided tests

$$H_0 : \mu \leq \mu_0 \quad \text{versus} \quad H_0 : \mu > \mu_0$$

Type II error probability at $\mu = \mu_a$:

$$\beta = P\left(z < z_\alpha - \frac{|\mu_a - \mu_0|}{\sigma / \sqrt{n}} \right)$$

Power (at $\mu = \mu_a$)

$$1 - \beta = P\left(z > z_\alpha - \frac{|\mu_a - \mu_0|}{\sigma / \sqrt{n}} \right)$$

$$H_0 : \mu \geq \mu_0 \quad \text{versus} \quad H_0 : \mu < \mu_0$$

Type II error probability at $\mu = \mu_a$:

$$\beta = P\left(z < z_\alpha - \frac{|\mu_a - \mu_0|}{\sigma / \sqrt{n}} \right)$$

Power (at $\mu = \mu_a$)

$$1 - \beta = P\left(z > z_\alpha - \frac{|\mu_a - \mu_0|}{\sigma / \sqrt{n}} \right)$$

Sample size required to have power = 1- β for $\Delta = |\mu_a - \mu_0|$

$$n = \frac{\sigma^2 \left(z_\alpha + z_\beta \right)^2}{\Delta^2}$$

Example

Ex. Egg weights: H_0: ≤ 65 H_a: μ > 65 (α = .05)

$$z_\alpha = 1.65 \quad \text{Reject if } z = \frac{\bar{y}-\mu_0}{\sigma_{\bar{y}}} = \frac{\bar{y}-\mu_0}{\sigma/\sqrt{n}} > 1.65$$

$$z = \frac{65.4-65}{.704} = \frac{.4}{.704} = .57 < 1.65 \Rightarrow \text{Do not reject } H_0$$

Power (at $\mu_a = 66$) = $P\left(z > z_\alpha - |65-66|/.704\right)$

= $P(z > 1.65 - 1.42)$ = P(z > 0.23) = .4090

Level of significance of statistical test

Consider tossing coin 10 times. Five heads and five tails are most probable event if coin is fair; i.e. if P(H) = .5. Six heads or seven heads would not cause doubt of the hypothesis that coin is fair. Eight or nine heads might stir discomfort. Ten heads would cause disbelief since the probability of ten heads is very small, 1/1024 = .00098. The distribution of the number of heads is binomial.

# Heads	6	7	8	8	10
Probability	.205	.117	.044	.0098	.00098

y = # heads

$P(y \geq 10) = .00098$

$P(y \geq 9) = .0098 + .00098 = .01078$

$P(y \geq 8) = .044 + .01078 = .0548$

$P(y \geq 7) = .117 + .0548 = .1718$

$P(y \geq 6) = .205 + .1718 = .3768$

As the number of heads increases, so does the doubt about fairness of the coin. This is because the small probability of a large number of heads makes fairness seem incredible. The degree of incredibility increases with the number of heads. We have "doubt" with 8 heads, "serious doubt" with 9 heads, and "disbelief" with 10 heads. This illustrates the principle of level of significance of statistical tests.

Level of significance of statistical test

Consider the one-sided hypothesis testing problem:

H_0: P(H) = .5 H_a: P(H) > .5

Test Statistic: y = # heads

Incredible set: large values of y

P-value (significance level) = P(y as extreme as observed)

So the P-value of 10 heads is $P(y \geq 10) = .00098$, the P-value of 9 heads is $P(y \geq 9) = .01078$, etc.

Level of significance of statistical test

We are not faced with making decisions in most research applications of statistical tests. Instead, we want to present measure of the evidence against the null hypothesis. Consider the egg weight problem.

$$H_0 : \mu = 65 \text{ versus } H_0 : \mu > 65$$

Test statistic: $z = (\overline{y} - 65)/(\sigma/\sqrt{n}) = .57$.

Instead of concluding "Reject H_0" or "Do not Reject H_0" we can report a *measure of incredibility* of H_0 as the probability of obtaining a value of z as large as the one observed if H_0 is true.

$$P(z \geq .57) = .284$$

The number .284 is the *significance probability*, which is often called the *p-value* of the test.

For two-sided test $H_0 : \mu = 65$ versus $H_a : \mu \neq 65$ the "significant" values of z are either large positively or negatively.

Then $p = P(|z| \geq .57) = 2(.284) = .568$

Inference about a Mean when σ^2 is Unknown

Formulas you have seen for test of hypothesis and confidence intervals for the mean of a normal distribution are based on the z distribution. These formulas entail the population variance σ^2.

In most applications, the population variance is not known and cannot be used in the formulas. You must substitute the sample variance s^2 in place of σ^2. If the sample size n is 30 or larger, you may proceed as usual using the values from the z distribution. But when $n<30$, use the t distribution rather than the z distribution.

Inference about a Mean when σ^2 is Unknown

Test of Hypothesis:

To test the null hypothesis $H_0: \mu = \mu_0$, use the test statistic

$$t = \frac{(\bar{y} - \mu_0)}{s / \sqrt{n}}$$

When the null hypothesis is true, this statistic has a t distribution with n-1degrees of freedom is, where n is the number of observations. Compute significance probabilities for the test using the t distribution.

Inference about a Mean when σ^2 is Unknown

Example: Following are monthly book sales for 12 months. The publisher projected average monthly sales of 2400. Is there statistical evidence in the data that sales exceed projection?

```
2277 2427 3206 2518 2251 2246 2268 2865 2915 2888 2047 2694
```

We wish to test the null hypothesis H_0: μ=2400 versus the alternative H_0: μ>2400. Since n=12 is less than 30 and we do not know the value of σ, we use the t statistic.

The mean, standard deviation and standard error of the mean are

$$\bar{y} = 2550.2 \quad s^2 = 127527 \quad s = 357.1 \quad s_{\bar{y}} = 103.1$$

The test statistic is

$$t = \frac{(\bar{y} - \mu_0)}{s / \sqrt{n}} = \frac{(\bar{y} - 2400)}{357.1 / \sqrt{12}} = \frac{150.2}{103.1} = 1.46$$

with n-1=11 degrees of freedom. The one-sided p-value is 0.087.

Inference about a Mean when σ^2 is Unknown

Finding a p-value from the t distribution:

Tables for the t distribution are not as complete as for the normal distribution because there is a different t distribution for each number of degrees of freedom. Complete tables would require at least a page for each number of degrees of freedom. Therefore, you cannot obtain exact p-values from t tables.

For example, the t statistic for the book sales data was $t = 1.46$ with 11 degrees of freedom. This number falls between $t_{.10} = 1.36$ and $t_{.05} = 1.79$. Therefore, the p-value is between .05 and .10. It is often adequate to report $p < .10$, or $.05 < p < .10$.

If you need a more precise p-value, you use a computer program such as SAS or Excel to obtain it. The p-value of 0.087 was obtained from Excel.

Inference about a Mean when σ^2 is Unknown

Alternatively, you can interpolate linearly from the tables, as you might have learned to do in high school trigonometry:

Compute the ratio (1.46 – 1.36)/(1.79 – 1.36) = .23. Then compute

$$\text{p-value} \approx .10 - (.23)(.10 - .05) = .10 - .012 = .088$$

This is a very good approximation the "exact" p-value of 0.087.

Note: The procedure illustrated is for an upper one-sided test. It can be adapted for a lower one-sided or a two-sided test.

Inference about a Mean when σ^2 is Unknown

Comparison of critical values of the t distribution with normal distribution for selected α and degrees of freedom.

			df = 60		df = 15	
	z_α	$z_{\alpha/2}$	t_α	$t_{\alpha/2}$	t_α	$t_{\alpha/2}$
$\alpha = .05$	1.65	1.96	1.67	2.00	1.753	2.131
$\alpha = .01$	2.31	2.57	2.39	2.66	2.60	2.95

Inference about a Mean when σ^2 is Unknown

Confidence Interval:

A 100(1-α) confidence interval for μ is given by

$$(\bar{y} - t_{\alpha/2} s/\sqrt{n}, \bar{y} + t_{\alpha/2} s/\sqrt{n})$$

The 95% confidence interval for the mean book sales is

(2550.2-2.20·(103.1), 2550.2+2.20·(103.1)) = (2343.4, 2757.0).

The number 2.20 is $t_{.025}$.