Math 1710 Class 26

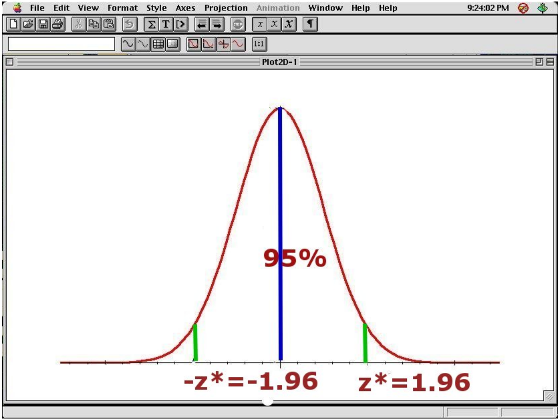
Hypothesis Testing

Dr. Back

Oct. 28, 2009

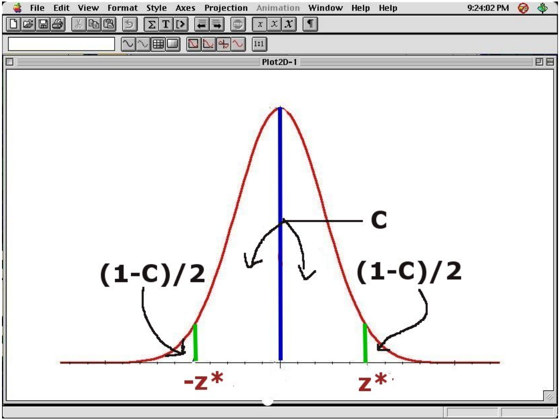
CI for p

Critical Value z^*. $(P(Z < z^*) = .975)$



CI for p

Critical Value z^* for a level C CI



CI for p

$$\hat{p} \pm z^* SE(\hat{p})$$

$$\text{where } SE(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n}}.$$

CI for p

$$\hat{p} \pm z^* SE(\hat{p})$$

$$\text{where } SE(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n}}.$$

Poll size $n = 400$, 144 say yes, $\hat{p} = .36$

CI for p

$$\hat{p} \pm z^* SE(\hat{p})$$

$$\text{where } SE(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n}}.$$

Poll size $n = 400$, 144 say yes, $\hat{p} = .36$

$$SE(\hat{p}) = \sqrt{\frac{.36 \cdot .64}{400}} = .024.$$

CI for p

$$\hat{p} \pm z^* SE(\hat{p})$$

$$\text{where } SE(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n}}.$$

Poll size $n = 400$, 144 say yes, $\hat{p} = .36$

$$SE(\hat{p}) = \sqrt{\frac{.36 \cdot .64}{400}} = .024.$$

Sampling distribution of $\hat{p}$ is $N(p, .024)$ approximately.

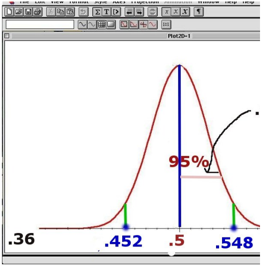
CI for p

$$\hat{p} \pm z^* SE(\hat{p})$$

$$\text{where } SE(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n}}.$$

Poll size $n = 400$, 144 say yes, $\hat{p} = .36$

This shows $p = .5$ should not be in the CI:



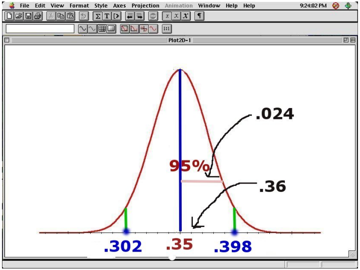
CI for p

$$\hat{p} \pm z^* SE(\hat{p})$$

$$\text{where } SE(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n}}.$$

Poll size $n = 400$, 144 say yes, $\hat{p} = .36$

This shows $p = .35$ should be in the CI:



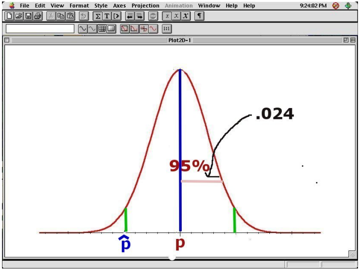
CI for p

$$\hat{p} \pm z^* SE(\hat{p})$$

$$\text{where } SE(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n}}.$$

Poll size $n = 400$, 144 say yes, $\hat{p} = .36$

Picture when p is at the upper limit of the CI:



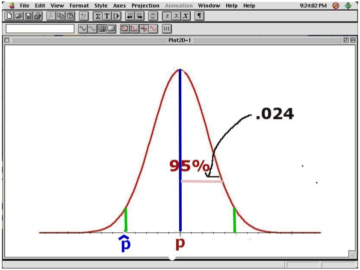
CI for p

$$\hat{p} \pm z^* SE(\hat{p})$$

$$\text{where } SE(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n}}.$$

Poll size $n = 400$, 144 say yes, $\hat{p} = .36$

Picture when p is at the upper limit of the CI:



The upper limit is $\hat{p} + z^* SE(\hat{p}) = .36 + 1.96 \cdot .024 = .407.$

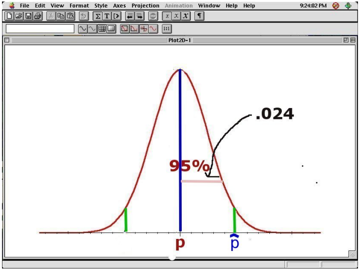
CI for p

$$\hat{p} \pm z^* SE(\hat{p})$$

$$\text{where } SE(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n}}.$$

Poll size $n = 400$, 144 say yes, $\hat{p} = .36$

Picture when p is at the lower limit of the CI:



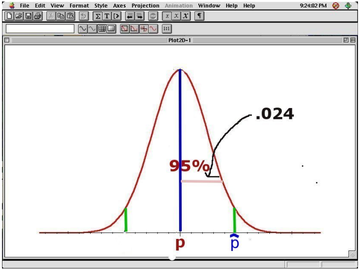
CI for p

$$\hat{p} \pm z^* SE(\hat{p})$$

$$\text{where } SE(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n}}.$$

Poll size $n = 400$, 144 say yes, $\hat{p} = .36$

Picture when p is at the lower limit of the CI:



The lower limit is $\hat{p} - z^* SE(\hat{p}) = .36 - 1.96 \cdot .024 = .313.$

CI Interpretation

Suppose we use a sample of size n to determine a 95% CI (a,b) for a parameter p.

CI Interpretation

If repeated with a large number of different samples of size n, (maybe different p's in each case as well) the defining property of CI's is that approximately 95% of CI's using samples of size n will contain the true parameter(s) p.

CI Interpretation

But the convention in elementary statistics is to view the statement "There is a 95% chance the CI contains p" to be false. That's because by the time the CI is computed, p already has a value and the statement p is in the CI is simply either true or false. No probability between 0 and 1.

CI Interpretation

But it is correct to say that 95% of the time, a 95% CI contains the true parameter.

CI Interpretation

Colloquially: A 95% CI uses a method which works 95% of the time.

CI Interpretation

works $\longleftrightarrow$ covers the true parameter

Conditions for Prop Tests

- plausible independence
- random sampling
- 10% condition
- success/failure

What Happens if Not Satisfied:

Conditions for Prop Tests

- plausible independence
- random sampling
- 10% condition
- success/failure

What Happens if Not Satisfied:

random sampling -
could be critical;
might be ok if "representative"
representative hard/impossible to define

Conditions for Prop Tests

- plausible independence
- random sampling
- 10% condition
- success/failure

What Happens if Not Satisfied:

plausible independence -
could be critical
sometimes just a working hypothesis

Conditions for Prop Tests

- plausible independence
- random sampling
- 10% condition
- success/failure

What Happens if Not Satisfied:

10% condition -
results in overestimation of samp. dist. st dev
gradual breakdown in formulas, not method

Conditions for Prop Tests

- plausible independence
- random sampling
- 10% condition
- success/failure

What Happens if Not Satisfied:

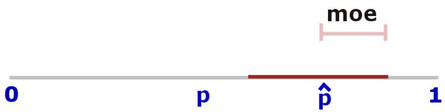
success/failure -
progressive reduction of accuracy
accuracy varies regardless for smaller values of n

Width of CI's

The margin of error in a CI is half its width:

Width of CI's

$$CI(red): \quad \hat{p} \pm z^* SE(\hat{p})$$
$$MOE(pink): \quad z^* SE(\hat{p})$$



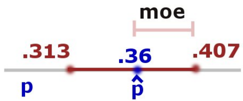
where $SE(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n}}$.

Width of CI's

Previous 95% CI Newspaper Poll Example

$CI(red): \quad \hat{p} \pm z^* SE(\hat{p}) \qquad .36 \pm .047$

$MOE(pink): \quad z^* SE(\hat{p}) \qquad 1.96 \cdot .024 = .047$



where $SE(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n}} = .024.$

Width of CI's

How big should a CI you are 100% certain of be?

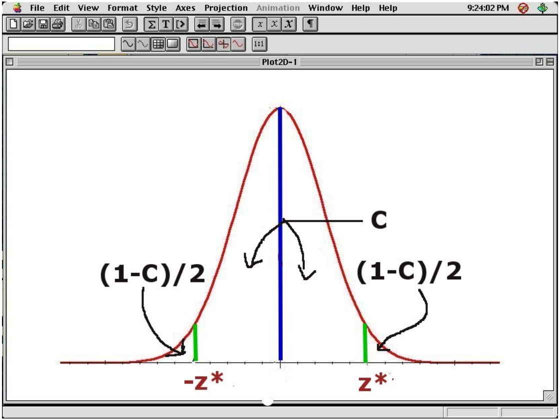
Width of CI's

How big should a CI you are 100% certain of be?

Answer: A 100% CI would be $(0, 1)$.

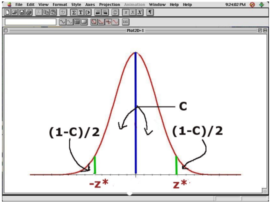
Width of CI's

Margins of error increase as the level of confidence increases.



Width of CI's

But not linearly.



Level of Confidence C	z^*
80%	1.282
90%	1.645
95%	1.96
98%	2.326
99%	2.576

Width of CI's

These numbers could come from table Z

$$\text{e.g. } z^* = 1.28 \;\; C = 80\%$$

z	Second decimal place in z									
	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319

Width of CI's

These numbers could come from table Z

e.g. $z^* = 1.28 \;\; C = 80\%$

	Second decimal place in z									
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319

Width of CI's

These numbers actually come from the bottom of table T

250	1.285	1.651	1.969	2.341	2.596
400	1.284	1.649	1.966	2.336	2.588
1000	1.282	1.646	1.962	2.330	2.581
∞	1.282	1.645	1.960	2.326	2.576
confidence levels	80%	90%	95%	98%	99%

Smoke Detectors

CPSC '96: 90% of American homes have at least one smoke detector. After a public safety campaign, a city observes that 376 out of 400 randomly selected homes have a detector.

Smoke Detectors

CPSC '96: 90% of American homes have at least one smoke detector. After a public safety campaign, a city observes that 376 out of 400 randomly selected homes have a detector.
Is this strong evidence the local rate is greater than the national rate?

Smoke Detectors

CPSC '96: 90% of American homes have at least one smoke detector. After a public safety campaign, a city observes that 376 out of 400 randomly selected homes have a detector. Is this strong evidence the local rate is greater than the national rate?

$$\text{So } \hat{p} = \frac{376}{400} = .94.$$

Logic of Hypothesis Testing

Null Hypothesis H_0 - retained unless disproven.

Logic of Hypothesis Testing

Null Hypothesis H_0 - retained unless disproven.
Alternative Hypothesis H_A- Only thing which possibly can be proven in the HT.

Logic of Hypothesis Testing

Null Hypothesis H_0 - retained unless disproven.
Alternative Hypothesis H_A- Only thing which possibly can be proven in the HT.
Procedure:

1. ...
2. ...
3. Consider the sampling distribution of $\hat{p}$ which would hold if H_0 were true.
4. Retain H_0 if $\hat{p}$ is reasonably consistent with this sampling distribution. Otherwise reject H_0.
5. ...

Logic of Hypothesis Testing

Procedure:

1. ...
2. ...
3. Consider the sampling distribution of $\hat{p}$ which would hold if H_0 were true.
4. Retain H_0 if $\hat{p}$ is reasonably consistent with this sampling distribution. Otherwise reject H_0.
5. ...

Three ways to carry out step 4:

1. Calculate a Z-statistic and determine a p-value.
2. Calculate a Z-statistic and compare with a critical value z^*.
3. Use a confidence interval.

Hypothesis Testing Vocabulary

- Null Hypothesis H
- Alternative Hypothesis H
  (One Sided vs Two Sided)
- $SD(\hat{p})$ vs. $SE(\hat{p})$
- Z-statistic
- Tail Probability
- P-value
- Significance Level
- Retain Null H vs. Accept Alt H (equiv. reject Null H)

CPSC Example

Let p be the true proportion of smoke detectors homes in our city have.

CPSC Example

Let p be the true proportion of smoke detectors homes in our city have.

$H_0 : p = .9, \quad H_A : p > .9$

CPSC Example

Let p be the true proportion of smoke detectors homes in our city have.
$H_0 : p = .9, \quad H_A : p > .9$

$$SD(\hat{p}) = \sqrt{\frac{.9 \cdot .1}{400}} = .015.$$

CPSC Example

Let p be the true proportion of smoke detectors homes in our city have.

$H_0 : p = .9, \quad H_A : p > .9$

$$SD(\hat{p}) = \sqrt{\frac{.9 \cdot .1}{400}} = .015.$$

z-statistic

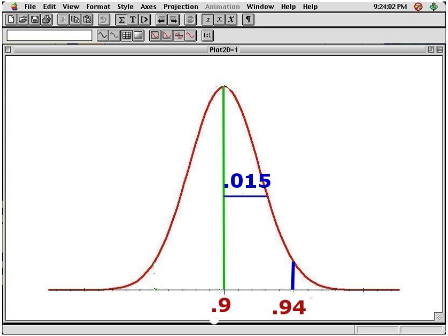
$$z = \frac{.94 - .9}{.015} = 2.67.$$

CPSC Example

z-statistic

$$z = \frac{.94 - .9}{.015} = 2.67.$$

z-statistic is the z-score of $\hat{p}$ on the samp dist centered at the hypothesized value of p

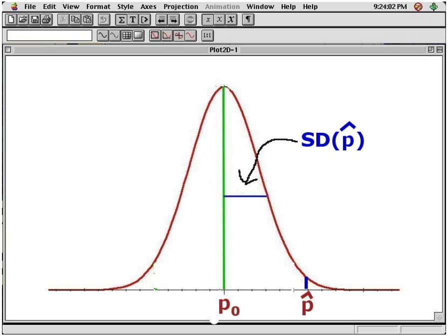


CPSC Example

z-statistic

$$z = \frac{.94 - .9}{.015} = 2.67.$$

z-statistic is the z-score of $\hat{p}$ on the samp dist centered at the hypothesized value p_0 of p



CPSC Example

z-statistic

$$z = \frac{.94 - .9}{.015} = 2.67.$$

Two primary ways to continue:

CPSC Example

z-statistic

$$z = \frac{.94 - .9}{.015} = 2.67.$$

Two primary ways to continue:
Method 1: Tail Prob $= P(\hat{p} > .94) = P(Z > 2.67) = .0038$
Since the test is 1-sided, P-value=tail probability=.0038.

CPSC Example

z-statistic

$$z = \frac{.94 - .9}{.015} = 2.67.$$

Two primary ways to continue:
Method 1: Tail Prob $= P(\hat{p} > .94) = P(Z > 2.67) = .0038$
Since the test is 1-sided, P-value=tail probability=.0038.
This is small, so conclusion is we reject H_0.

CPSC Example

z-statistic

$$z = \frac{.94 - .9}{.015} = 2.67.$$

Two primary ways to continue:

Method 2: Say significance level is $\alpha = .05$.

Since our z-statistic of 2.67 is more extreme than $z^* = 1.645$ (and supports H_a), we reject H_a at $\alpha = .05$.

CPSC Example

z-statistic

$$z = \frac{.94 - .9}{.015} = 2.67.$$

Two primary ways to continue:

N(0,1) Picture behind $z^* = 1.645$