Analysis of Economics Data
Chapter 4: Statistical Inference for the Mean

© A. Colin Cameron
Univ. of Calif. Davis

November 2022

CHAPTER 4: Statistical Inference

- Extrapolate from sample mean $\bar{x}$ to population mean μ.
- Given the sample, confidence intervals give a range of values that μ is likely to full into.
- Hypothesis tests are used to determine whether or not a specified value or range of values of μ is plausible, given the sample.
- While we focus on μ, the methods generalize to inference on other parameters.

Outline

- Datasets: EARNINGS, GASPRICE, EARNINGSMALE, REALGDPPC.

4.1 Example: Mean Annual Earnings

- Sample of 171 female full-time workers aged 30 in 2010.
- Descriptive statistics obtained using Stata `summarize` command

```
. summarize earnings
```

Variable	Obs	Mean	Std. Dev.	Min	Max
earnings	171	41412.69	25527.05	1050	172000

- Key statistics:
  - Mean: sample mean $\bar{x}$
  - Std. Dev.: standard error s measures the precision of $\bar{x}$ as an estimate of μ.
- The next slides present methods for statistical inference on μ that are explained in detail in the remainder of the chapter.

95% Confidence Interval for the Mean

- A 95% confidence interval for a parameter is a range of likely values that the parameter lies in with 95% confidence.
- 95% Confidence interval for μ obtained using Stata `mean` command.

```
. mean earnings

Mean estimation                   Number of obs   =        171

--------------------------------------------------------------
             |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
    earnings |   41412.69   1952.103      37559.21    45266.17
--------------------------------------------------------------
```

- Key statistics:
  - Mean: sample mean $\bar{x}$ is the estimate of μ
  - Std. Err: standard error measures the precision of $\bar{x}$ as an estimate of μ
    - ★ this equals $s/\sqrt{n} = 25527.05/\sqrt{171} = 1952.1$.

95% Confidence Interval Calculation

- In general a confidence interval is

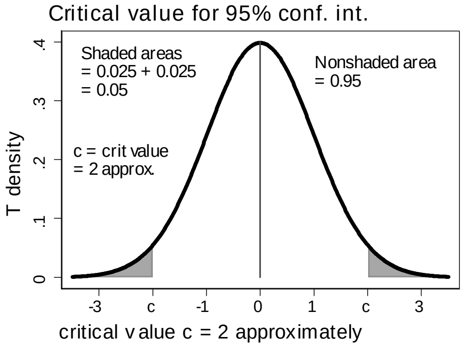
$$\text{estimate} \pm \text{ critical value} \times \text{standard error}$$

- Here we consider the population mean μ.
- The estimate is $\bar{x} = 41412.69$
- The standard error measures the precision of $\bar{x}$ as an estimate of μ
  - $se(\bar{x}) = s/\sqrt{n} = 25527.05/\sqrt{171} = 1952.1.$
- The 95% critical value is approximately 2
  - more precisely here $c = 1.974$ as $\Pr[|T_{170}| \leq 1.974] = 0.95.$
- The 95% confidence interval is then

$$\bar{x} \pm c \times se(\bar{x}) = 41412.69 \pm 1.974 \times 1952.1 = (37559,\ 45266).$$

Critical Value for the Confidence Interval

- For μ use the T distribution with $n-1$ degrees of freedom
  - very similar to standard normal distribution except with fatter tails.
- Let T_{n-1} denoted a random variable that is $T(n-1)$ distributed.
- The critical value c for a 95% conf. interval is that value for which
  - the probability that $|T_{n-1}| \leq c = 0.95$
  - equivalently the probability that $T_{n-1} \geq c = 0.05/2 = 0.025$.



Hypothesis test on the Mean

- Hypothesis test using Stata `ttest` command
  - as illustrative example test whether or not $\mu = 40,000$.

```
. ttest earnings = 40000

One-sample t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
earnings |     171    41412.69    1952.103    25527.05    37559.21    45266.17
------------------------------------------------------------------------------
    mean = mean(earnings)                                         t =   0.7237
Ho: mean = 40000                                 degrees of freedom =      170

    Ha: mean < 40000           Ha: mean != 40000             Ha: mean > 40000
 Pr(T < t) = 0.7649       Pr(|T| > |t|) = 0.4703          Pr(T > t) = 0.2351
```

- We test $H_0 : \mu = 40000$ against $H_a : \mu \neq 40000$.
- The test statistic is $t = 0.7237$.
- The p-value is 0.4703 (as we test against $H_a : \mu \neq 40000$).
- Since $p > 0.05$ we do not reject $H_0 : \mu = 40000$ at level 0.05.

Hypothesis test calculation

- In general a t test statistic is

$$t = \frac{\text{estimate} \ - \ \text{hypothesized value}}{\text{standard error}} \ .$$

- Here

$$t = \frac{\bar{x} - \mu_0}{se(\bar{x})} = \frac{41412.69 - 40000}{1952.1} = 0.7237.$$

- The p-value is the probability of observing a value at least as large as this in absolute value.
- Here p equals the probability that $|T_{170}| \geq 0.7237 = 0.4703$.
- Since this probability exceeds 0.05 we do not reject H_0.

4.2 t Statistic and t distribution

- Estimate μ using $\bar{x}$ which is the sample value of draw of the random variable $\bar{X}$
- So far we have $E[\bar{X}] = \mu$ and $Var[\bar{X}] = \sigma^2/n$ for a simple random sample.
- For confidence intervals and hypothesis tests on μ we need a distribution
  - under certain assumptions $\bar{X}$ is normally distributed
  - but with variance that depends on the unknown σ^2
  - we replace σ^2 by the estimate s^2
  - this leads to use of the t-statistic and the t distribution
    - similar to the standard normal but with fatter tails.

Normal Distribution and the Central Limit Theorem

- We assume a **simple random sample** where
  - **A.** X_i has common mean μ : $\mathrm{E}[X_i] = \mu$ for all i.
  - **B.** X_i has common variance σ^2 : $\mathrm{Var}[X_i] = \sigma^2$ for all i.
  - **C.** Statistically independence: X_i is statistically independent of $X_j, i \neq j$.
- Then $\bar{X} \sim (\mu, \sigma^2/n)$, i.e. $\bar{X}$ has mean μ and variance σ^2/n.
- Under these assumptions the standardized variable $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim (0,1)$.
- The central limit theorem (a remarkable result) states that if additionally the sample size is large Z is normally distributed

$$Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1) \text{ as } n \rightarrow \infty.$$

The t-statistic

- Now replace the unknown σ^2 by an estimator $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$.

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}.$$

- The distribution for T is complicated. The standard approximation is T has the t distribution with $(n-1)$ degrees of freedom

$$T \sim T(n-1)$$

- Comments
  - different degrees of freedom correspond to different t distributions
  - the term degrees of freedom is used because $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ implies that only $(n-1)$ terms in the sum are free to vary
  - $T \sim T(n-1)$ exactly in the very special case that $X_i s$ are normally distributed
  - otherwise T is not $T(n-1)$ exactly but is the standard approximation.

The t-statistic (continued)

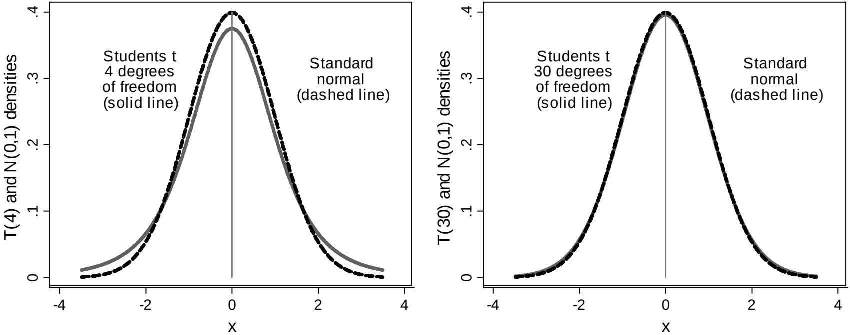
- In summary, inference on μ is based on the sample t-**statistic** is

$$t=\frac{\bar{x}-\mu}{se(\bar{x})}=\frac{\bar{x}-\mu}{s/\sqrt{n}},$$

  - $\bar{x}$ is the sample mean
  - $se(\bar{x})$ is the standard error of $\bar{x}$
  - s is the sample standard deviation.
- The statistic t is viewed as a realization of the $T(n-1)$ distribution.

The t Distribution

- t **distribution** has probability density function that is bell-shaped
  - $\Pr[a < T < b]$ is the area under the curve between a and b
- The t distribution has fatter tails than the standard normal.
- T_v denotes a random variable that has the $T(v)$ distribution.
- Different values of v correspond to different T distributions
  - t_∞ is the same as $N(0,1)$.



Probabilities for the t Distribution

- **Probabilities** are the area under the t probability density function.
  - e.g. $\Pr[a < T < b]$ is the area under the curve from a to b
- Computing these probabilities requires a computer.
- The Stata function `ttail(v,t)` gives $\Pr[T_v > t]$
  - e.g. $\Pr[T_{170} > 0.724] =$ `ttail(170,0.724)`$= 0.235$.
- The R function `1-pt(t,v)` gives $\Pr[T_v > t]$
  - e.g. $\Pr[T_{170} > 0.724] =$ `1-pt(0.724,170)`$= 0.235$.

Inverse Probabilities for the t Distribution

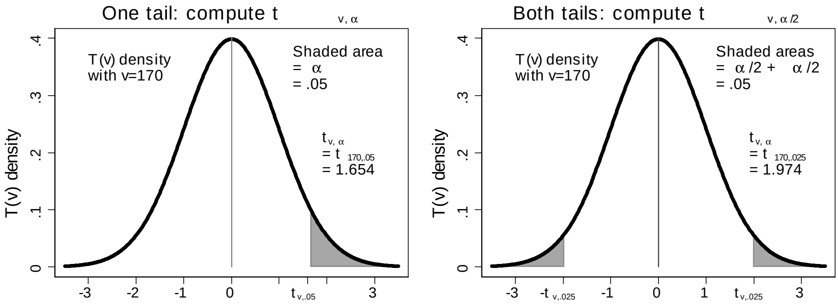
- For confidence intervals we need to find the inverse probability
  - called a critical value.
- Definition: the **inverse probability** or **critical value** $c = t_{v,\alpha}$ is that value such that the probability that a $T(v)$ distributed random variable exceeds $t_{v,\alpha}$ equals α.

$$\Pr[T_v > t_{v,\alpha}] = \alpha.$$

  - i.e. the area in the right tail beyond $t_{v,\alpha}$ equals α.
- Example: $\Pr[T_{170} > 1.654] = 0.05$ so $c = t_{170,.05} = 1.654$.
- The Stata function `invttail(v,a)` gives a such that $\Pr[T_v > t] = \text{a}$
  - e.g. $c = t_{170,.05}$ =`invttail(170,.05)`$= 1.654$.
- The R function is `qt(1-a,v)` e.g. `qt(0.95,170)`$= 1.654$.

Inverse probabilities (continued)

- Left panel: $\Pr[T_{170} > 1.654] = 0.05$, so $t_{170,.05} = 1.654$.
- Right panel: $\Pr[-1.974 < T_{170} < 1.974] = 0.05$
  - using $\Pr[T_{170} > 1.974] = 0.025$ and $t_{170,.025} = 1.974$.



4.3 Confidence Intervals

- For simplicity focus on 95% confidence intervals.
- A **95 percent confidence interval for the population mean** is

$$\bar{x} \pm t_{n-1,.025} \times se(\bar{x}),$$

  - $\bar{x}$ is the sample mean
  - $t_{n-1,.025}$ is exceeded by a $T(n-1)$ random variable with probability 0.025
  - $se(\bar{x}) = s/\sqrt{n}$ is the standard error of the sample mean.
- The area in the tails is $0.025 + 0.025 = 0.05$
  - leaving area 0.95 in the middle
  - hence a 95% confidence interval.

Example: Mean Annual Earnings

- Here $\bar{x} = 41413$, $se(\bar{x}) = s/\sqrt{n} = 1952$, $n = 171$, and $t_{170,.025} = 1.974$.
- A 95% confidence interval (CI) is

$$\begin{aligned} \bar{x} \pm t_{n-1,\alpha/2} \times \left(s/\sqrt{n}\right) &= 41413 \pm 1.974 \times 1952 \\ &= 41413 \pm 3853 \\ &= (37560,\ 45266). \end{aligned}$$

- A 95% confidence interval for population mean earnings of thirty year-old female full-time workers is
  - ($37,560, $45,266)
  - this was the result obtained earlier using the Stata `mean` command.

Derivation of a 95% Confidence Intervals

- We derive a 95% confidence interval from first principles.
- For simplicity consider a sample with $n = 61$, in which case $n - 1 = 60$ and $t_{60,.025} = 2.0003$. Thus

$$\Pr[-2.0003 < T_{60} < 2.0003] = 0.95.$$

- Round to $\Pr[-2 < T < 2] = 0.95$ and substituting $T = \frac{\bar{X}-\mu}{S/\sqrt{n}}$ yields

$$\Pr\left[-2 < \frac{\bar{X}-\mu}{S/\sqrt{n}} < 2\right] = 0.95.$$

- Convert to an interval that is centered on μ as follows

$$\begin{aligned}
\Pr\left[-2 < \frac{\bar{X}-\mu}{S/\sqrt{n}} < 2\right] &= 0.95 & \\
\Pr\left[-2S/\sqrt{n} < \bar{X}-\mu < 2S/\sqrt{n}\right] &= 0.95 & \text{times } S/\sqrt{n} \\
\Pr\left[-\bar{X}-2S/\sqrt{n} < -\mu < -\bar{X}+2S/\sqrt{n}\right] &= 0.95 & \text{subtract } \bar{X} \\
\Pr\left[\bar{X}+2S/\sqrt{n} > \mu > \bar{X}-2S/\sqrt{n}\right] &= 0.95 & \text{times } -1.
\end{aligned}$$

Derivation (continued)

- Re-ordering the final inequality yields

$$\Pr\left[\bar{X} - 2 \times S/\sqrt{n} < \mu < \bar{X} + 2S/\sqrt{n}\right] = 0.95.$$

- Replace random variables by their observed values
  - the interval $(\bar{x} - 2 \times s/\sqrt{n}, \ \bar{x} + 2 \times s/\sqrt{n})$ is called a 95% confidence interval for μ.
- More generally with sample size n the critical value is $t_{n-1,.025}$.
- A 95% confidence interval is $(\bar{x} - t_{n-1,.025} \times se(\bar{x}),$ $\bar{x} + t_{n-1,.025} \times se(\bar{x}))$.
- This is the confidence interval formula given earlier.

What Level of Confidence?

- Ideally narrow confidence intervals with high level of confidence.
- But trade-off: more confidence implies wider interval
  - e.g. 100% confidence is μ in $(-\infty, \infty)$.
- What value of confidence should we use?
  - no best value in general
  - common to use a 95% confidence interval.
- A **$100(1-\alpha)\%$ percent confidence interval for the population mean** is

$$\bar{x} \pm t_{n-1,a/2} \times \left(s/\sqrt{n}\right).$$

  - $\alpha = 0.05$ (so $\alpha/2 = 0.025$) gives a 95% confidence interval as $100 \times (1-0.05) = 95$.
  - next most common are 90% $(\alpha = 0.10)$ and 99% $(\alpha = 0.01)$ confidence intervals

Critical t values

- Table presents $t_{v,\alpha/2}$ for various confidence levels (α) and $v = n - 1$.
- The 95% confidence intervals critical values are bolded

Confidence Level	$100(1-\alpha)$	90%	**95%**	99%
Area in both tails	α	0.10	**0.05**	0.01
Area in single tail	$\alpha/2$	0.05	**0.025**	0.005
t value for $v = 10$	$t_{10,\alpha/2}$	1.812	**2.228**	3.169
t value for $v = 30$	$t_{30,\alpha/2}$	1.697	**2.042**	2.750
t value for $v = 100$	$t_{100,\alpha/2}$	1.660	**1.980**	2.626
t value for $v = \infty$	$t_{\infty,\alpha/2}$	1.645	**1.960**	2.576
standard normal value	$z_{\alpha/2}$	1.645	**1.960**	2.576

- Note that $t_{v,.025} \simeq 2$ for $v > 30$.
- An approximate 95% confidence interval for μ is therefore a two-standard error interval
  - the sample mean plus or minus two standard errors.

Interpretation

- Interpretation of confidence intervals is conceptually difficult.
- The correct interpretation of a 95 percent confidence interval is that if constructed for each of an infinite number of samples then it will include μ 95% of the time
  - of course we only have one sample.
- 1880 Census example (we know $\mu = 24.13$) in Chapter 3
  - First sample of size 25: 95% confidence interval (17.99, 34.81)
  - Second sample: 95% CI (13.12, 25.54), and so on.
- For the particular 100 samples drawn
  - two samples had 95% confidence intervals that did not include μ
    - ★ 20^{th} sample had 95% interval (8.57, 23.90)
    - ★ 50^{th} sample had 95% interval (11.49, 21.45)
  - so here 98% of the samples had 95% confidence interval that included μ (versus theory 95%).

4.4 Two-Sided Hypothesis Tests

- A **two-sided test** or **two-tailed test** for the population mean is a test of the **null hypothesis**

$$H_0 : \mu = \mu^*$$

where μ^* is a specified value for μ, against the **alternative hypothesis**

$$H_a : \mu \neq \mu^*.$$

- In the next example $\mu^* = 40000$.
- Called two-sided as the alternative hypothesis includes both $\mu > \mu^*$ and $\mu < \mu^*$.
- We need to either reject H_0 or not reject H_0.

Significance Level of a Test

- A test either rejects or does not reject the null hypothesis.
- The decision made may be in error.
- A **type I error** occurs if H_0 is rejected when H_0 is true.
  - e.g. H_0 is person is innocent. A type I error is to reject H_0 and find the person guilty, when in fact the person was innocent.
- The **significance level** of a test, denoted α, is the pre-specified maximum probability of a type I error that will be tolerated.
- Often $\alpha = 0.05$. A 5% chance of making a type I error.

The t-test Statistic

- Obviously reject $H_0 : \mu = \mu^*$ if $\bar{x}$ is a long way from μ^*.
- Transform to $t = (\bar{x} - \mu^*)/se(\bar{x})$ as this has known distribution.
- Equivalently reject H_0 : if the t statistic is large in absolute value where

$$t = \frac{\bar{x} - \mu^*}{se(\bar{x})} = \frac{\bar{x} - \mu^*}{s/\sqrt{n}}.$$

- Example: Test whether or not population mean female earnings equal $\$40,000$.
- Here $H_0 : \mu = 40000$ and $n = 171$, $\bar{x} = 41412$, $s = 25527$, so $se(\bar{x}) = s/\sqrt{n} = 1952$

$$t = \frac{\bar{x} - \mu}{se(\bar{x})} = \frac{41412 - 40000}{1952} = 0.724.$$

- The t-statistic is a draw from the $T(170)$ distribution, since $n = 171$.

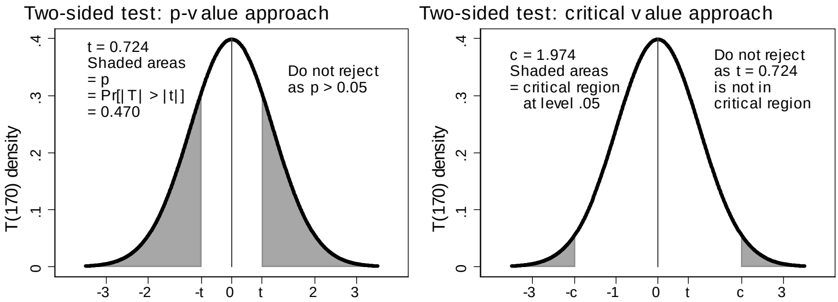
Rejection Using p-values

- How likely are we to obtain a draw from $T(170)$ that is $\geq |0.724|$?
- The *p*-**value** is the probability of observing a t-test statistic at least as large in absolute value as that obtained in the current sample.
- For a two-sided test of $H_0 : \mu = \mu^*$ against $H_a : \mu \neq \mu^*$ the p-value is

$$p = \Pr[|T_{n-1}| \geq |t|].$$

- H_0 is rejected at significance level α if $p < \alpha$, and is not rejected otherwise.
- Earnings example
  - $p = \Pr[|T_{170}| \geq 0.724] = 0.470$.
  - since $p > 0.05$ we do not reject H_0.

- Left panel: p-value
- Right panel: critical value



Rejection using Critical Regions

- Alternative equivalent method is the following
  - base rejection directly on the value of the t-statistic
  - requires table of critical values rather than computer for p-values.
- A **critical region** or **rejection region** is the range of values of t that would lead to rejection of H_0 at the specified significance level α.
- For a two-sided test of $H_0 : \mu = \mu^*$ against $H_a : \mu \neq \mu^*$, and for specified α, the **critical value** c is such that

$$c = t_{n-1,\alpha/2} \text{ (so equivalently } \Pr[|T_{n-1}| \geq c] = \alpha).$$

- H_0 is rejected at significance level α if $|t| > c$, and is not rejected otherwise.
- Earnings example:
  - if $\alpha = 0.05$ then $c = t_{170,0.025} = 1.974$.
  - do not reject H_0 since $t = 0.724$ and $|0.724| < 1.974$.
- The critical value is illustrated in right panel of the preceding figure.

Which Significance level?

- Decreasing the significance level α
  - decreases the area in the tails that defines the rejection region
  - makes it less likely that H_0 is rejected.
- It is most common to use $\alpha = 0.05$, called a test at the 5% significance level
  - then a type I error is made 1 in 20 times.
- This is a convention and in many applications other values of α may be warranted.
  - e.g. What if H_0 : no nuclear war? Then use $\alpha > 0.05$.
- Reporting p-values allows the reader to easily test using their own preferred value of α.
- Further discussion under test power.

Relationship to Confidence Intervals

- Two-sided tests can be implemented using confidence intervals.
- If the H_0 value μ^* falls inside the $100(1-\alpha)$ percent confidence interval then do not reject H_0 at level α.
- Otherwise reject H_0 at significance level α.

Summary

- A summary of the preceding example is the following.

Hypotheses	$H_0 : \mu = 40000 \quad H_a : \mu \neq 40000 \quad \alpha = 0.05$
Significance level	$\alpha = 0.05$
Sample data	$\bar{x} = 41412$, $s = 25527$, $n = 171$
Test statistic	$t = (41412 - 40000)/(25527/\sqrt{171}) = 0.724$
(1) p-value	$p = \Pr[\|T_{170}\| \geq \|0.724\|] = 0.470$
approach	Do not reject H_0 at level .05 as $p > .05$
(2) Critical value	$c = t_{170,.025} = 1.974$
approach	Do not reject H_0 at level .05 as $\|t\| < c$.

- The p-value and critical value approaches are alternative methods that lead to the same conclusion.

4.5 Hypothesis Testing Example 1: Gasoline Prices

- Test at $\alpha = .05$ claim that the price of regular gasoline in Yolo County is neither higher nor lower than the norm for California.
  - one day's data from a website that provides daily data on gas prices
  - average California price that day was $3.81
  - $H_0 : \mu = 3.81$ is tested against $H_a : \mu \neq 3.81$.
- $n = 32$, $\bar{x} = 3.6697$ and $s = 0.1510$.
- $t = (3.6697 - 3.81)/(0.1510/\sqrt{32}) = -5.256$.
- p value method: $p = \Pr[|T_{31}| > 5.256] = 0.000$
  - reject H_0 at level .05 since $p < .05$.
- Critical value method: $c = t_{31,.025} = 2.040$.
  - reject H_0 at level .05 since $|t| = 5.256 > c = 2.040$.
- Reject the claim that reject the claim that population mean Yolo County gas price equals the California state-average price.

Example 2: Male Earnings

- Test at $\alpha = .05$ the claim that population mean annual earnings for 30 year-old U.S. men with earnings in 2010 exceed \$50,000
  - claim that > 50000 is set up as the alternative hypothesis
  - $H_0 : \mu \leq 50000$ is tested against $H_a : \mu > 50000$.
- $n = 191$, $\bar{x} = 52353.93$ and $s = 65034.74$.
- $t = (52353.93 - 50000)/(65034.74/\sqrt{191}) = 0.5002$.
- p value method: $p = \Pr[T_{190} > 0.500] = 0.310$.
  - do not reject H_0 at level .05 since $p > .05$.
- Critical value method: $c = t_{190,.05} = 1.653$.
  - do not reject H_0 at level .05 since $t = 0.500 > c = 1.653$.
- Do not reject the claim that population mean earnings exceed \$50,000.

Example 3: Price Inflation

- Test at $\alpha = .05$ claim that U.S. real GDP per capita grew on average at 2.0% over the period 1960 to 2020
  - use year-to-year percentage changes in U.S. real GDP per capita.
  - $H_0 : \mu = 2.0$ tested against $H_a : \mu \neq 2.0$.
- $n = 241$, $\bar{x} = 1.9904$ and $s = 2.1781$.
- $t = (1.9904 - 2.0)/(2.1781/\sqrt{241}) = -0.068$.
- p value method: $p = \Pr[|T_{258}| > 0.0680] = 0.946$
  - do not reject H_0 at level .05 since $p < .05$.
- Critical value method: $c = t_{241,.025} = 1.970$
  - do not reject H_0 at level .05 since $|t| = 0.068 < c = 1.970$.
- Do not reject the claim that population mean growth was 2.0%.

4.6 One-sided Directional Hypothesis Tests

- An **upper one-tailed alternative test** is a test of $H_0 : \mu \leq \mu^*$ against $H_a : \mu > \mu^*$.
- A **lower one-tailed alternative test** is a test of $H_0 : \mu \geq \mu^*$ against $H_a : \mu < \mu^*$.
- For one-sided tests the statement being tested is specified to be the alternative hypothesis.
- And if a new theory is put forward to supplant an old, the new theory is specified to be the alternative hypothesis.
- Example: Test claim that population mean earnings exceed $40,000
    - test $H_0 : \mu \leq 40000$ against $H_a : \mu > 40000$.

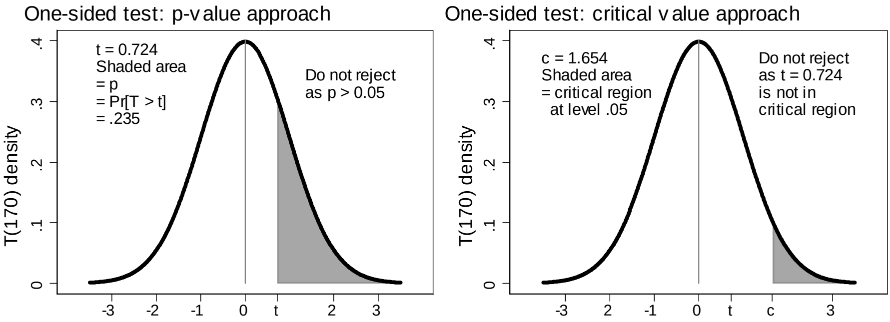
P-Values and Critical Regions

- Use the usual t-test statistic $t = (\bar{x} - \mu^*)/se(\bar{x})$.
- For an **upper one-tailed alternative** test
  - $p = \Pr[T_{n-1} \geq t]$ is p-**value**
  - $c = t_{n-1,\alpha}$ is **critical value** at significance level α
  - reject H_0 if $p < \alpha$ or, equivalently, if $t > c$.
- For a **lower one-tailed alternative** test
  - $p = \Pr[T_{n-1} \leq t]$ is p-**value**
  - $c = -t_{n-1,\alpha}$ is **critical value** at significance level α
  - H_0 if $p < \alpha$ or, equivalently, if $t < c$.

Example: Mean Annual Earnings

- Evaluate the claim that the population mean exceeds $40,000.
- Test of $H_0 : \mu \leq 40000$ against $H_a : \mu > 40000$
  - the claim is specified to be the alternative hypothesis
  - a detailed explanation is given next
  - and we reject if t is large and positive.
- From earlier $t = 0.724$.
- p value method: $p = \Pr[T_{170} \geq .724] = 0.235$
  - do not reject H_0 at level 0.05 since $p > 0.05$.
- Critical value method: $c = t_{170,.05} = 1.654$
  - do not reject H_0 at level 0.05 since $t = 0.724 < c = 1.654$.

- Left panel: p-value
- Right panel: critical value



Specifying the Null Hypothesis in One-sided Test

- Suppose claim is that population mean earnings exceed $40,000.
- Potential method 1: test $H_0 : \mu \leq 40000$ against $H_a : \mu > 40000$
  - Reject H_0 if $\bar{x}$ quite a bit higher than 40000. e.g. 43,000.
  - Then claim that $\mu > 40000$ is supported if $\bar{x} > 43000$.
- Potential method 2: test $H_0 : \mu \geq 40000$ against $H_a : \mu < 40000$
  - Reject H_0 if $\bar{x}$ quite a bit smaller than than 40000. e.g. 37,000.
  - So do not reject H_0 if $\bar{x} > 37000$.
  - Then claim that $\mu > 40000$ is supported if $\bar{x} > 37000$
  - Much more likely to accept the claim than with method 1.
- The statistics philosophy: need strong evidence to support a claim
  - the first specification is therefore used
  - **the statement being tested is specified to be the alternative hypothesis.**

4.7 Generalize Confidence Intervals and Hypothesis Tests

- Consider general case of an estimate of a parameter
  - with standard error the estimated standard deviation of the estimate
  - generalizes $\bar{x}$ is an estimate of μ with standard error $se(\bar{x})$.
- For the models and assumptions considered in this book

$$t = \frac{\text{estimate} - \text{parameter}}{\text{standard error}} \sim T(v) \text{ distribution}$$

where the degrees of freedom v vary with the setting.

- The $100(1-\alpha)\%$ **confidence interval** for the unknown parameter is

$$\text{estimate} \pm t_{v,\alpha/2} \times \text{standard error.}$$

- Most often use 95% confidence level and $t_{v,.025} \simeq 2$ for $v > 30$.
- So an approximate 95% CI is a **two-standard error interval**

$$\text{estimate} \pm 2 \times \text{standard error.}$$

- **Margin of error** in general is half the width of a confidence interval.
  - For 95% confidence intervals, since $t_{v,.025} \simeq 2$,

$$\text{Margin of error} \simeq 2 \times \text{Standard error.}$$

Generalization of Hypothesis Tests

- Two-sided test at significance level α of
  - H_0 : a parameter equals a hypothesized value against
  - H_a : that it does not.
- Calculate the t-**statistic**

$$t = \frac{\text{estimate} - \text{hypothesized parameter value}}{\text{standard error}}.$$

  - under H_0 t is the sample realization of a $T(v)$ random variable.
- Two-sided hypothesis test at significance level α:
  - p-**value approach**: reject H_0 if $p < \alpha$ where $p = \Pr[|T_v| > t]$
  - **critical value approach**: reject H_0 if $|t| > c$ where $c = t_{v,\alpha/2}$ satisfies $\Pr[T_v > t_{v,\alpha/2}] = \alpha$
  - the two methods lead to the same conclusion.

4.8 Proportions Data

- Consider proportion of respondents voting Democrat.
- Code data as $x_i = 1$ if vote Democrat and $x_i = 0$ if vote Republican
  - the sample mean $\bar{x}$ is the proportion voting Democrat.
  - the sample variance $s^2 = n\bar{x}(1-\bar{x})/(n-1)$
    - ★ in this special case of binary data.
- Example: 480 of 921 voters intend to vote Democrat (and 441 vote Republican)
  - $\bar{x} = (480 \times 1 + 440 \times 0)/921 = 0.5212$
  - $s^2 = 921 \times 0.5212 \times (1 - 0.5212)/920 = 0.2498.$

Inference for Proportions Data

- View each outcome as result of random variable

$$X = \begin{cases} 1 & \text{with probability } p & \text{if vote Democrat} \\ 0 & \text{with probability } 1-p & \text{if vote Republican} \end{cases}$$

- Then $\bar{X}$ has mean p and variance $\sigma^2/n = p(1-p)/n$.
- Can do analysis using earlier results with the usual standard error of $\bar{x}$
  - here $s^2/n = n\bar{x}(1-\bar{x})/(n-1) = \bar{x}(1-\bar{x})/(n-1)$
- But usually confidence intervals substitute $\bar{x}$ for p in $\sigma^2/n = p(1-p)/n$
  - so standard error of $\bar{x}$ is $\bar{x}(1-\bar{x})/n$
- And hypothesis tests of $H_0 : p = p^*$ also substitute for p and use

$$t = \frac{\bar{x} - p^*}{\sqrt{p^*(1-p^*)/n}}.$$

Key Stata Commands

```
use EARNINGSBOTH.DTA, clear
* Confidence interval
mean earnings
mean earnings, level(90)
* Hypothesis test
ttest earnings = 40
* Upper tail probability
display ttail(170,0.724)
* Critical value or inverse tail probability
display invttail(170,0.025)
```

Computing the p-value and Critical Value

- Example of computer commands to get p and c
  - for $t = t$, degrees of freedom v, and test at level α
- Two-sided tests
  - Stata: $p = 2{*}\text{ttail}(v, |t|)$ and $c = \text{invttail}(v, \alpha/2)$
  - R: $p = 2 * (1{-}\texttt{pt}(|t|, v))$ and $c = \text{qt}(1 - \alpha/2, v)$
  - Excel: $p = \text{TDIST}(|t|, v, 2)$ and $c = \text{TINV}(2\alpha, v)$

Some in-class Exercises

1. Suppose observations in a sample of size 25 have mean 200 and standard deviation of 100. Give the standard error of the sample mean.
2. Suppose $n = 100$, $\bar{x} = 500$ and $s = 400$. Provide an approximate 95% confidence interval for the population mean.
3. Suppose observations in a sample of size 100 have mean 300 and standard deviation of 90. Test the claim that the population mean equals 280 at the 5% significance level.