

EEP/IAS 118 - Introductory Applied
Econometrics, Lecture 6

Gregory Lane

July 2017

Hypothesis Testing: Intro

Last time how to construct/ interpret confidence intervals.
Sometimes we may want to go beyond just generating a range of probable values for our population parameter:

- Often we want to test a hypothesis that our parameter is a specific value
- For example: “vocational training schools are effective at increasing employment rates”
- We will use our sample of data to test whether some hypothesis we have about the true population is likely or not.

Hypothesis Testing Process

To actually test a hypothesis, we need to follow five steps:

- 1 State the **null** (H_0) and **alternative** (H_1) hypotheses
 - The null hypothesis is presumed true until we have evidence to reject it
 - Most common is testing that the population parameter is zero $H_0 : \mu = 0$ against the alternative $H_1 : \mu \neq 0$
 - The form of the alternative will determine whether you perform a one-sided or two-sided test
- 2 Choose and calculate a **test statistic** with a *known* distribution. The test statistic is just a function of our random sample (a transformation of the estimator)
 - The test statistic provides a measure of how far our sample estimator (\bar{X}) is from our hypothesized population value (μ) relative to the standard error of the estimator ($se(\bar{X})$)

Hypothesis Testing Process

- 3 Choose **significance level** (α) and find **critical value** (c)
 - Sig. level is typically 10%, 5%, or 1%. This value is the probability that we reject the null when the null is true
 - Critical value is found on the appropriate table (t , z , or F) for the chosen significance level and given degrees of freedom
 - Be careful, the critical value will change depending on whether the test is one-sided or two-sided!
- 4 Define our rejection rule. Reject H_0 if and only if:

$$|t| > |c|$$

- 5 Interpret. We either:
 - a. Reject the null
 - b. Fail to reject the null

We **NEVER** accept the null

Hypothesis Testing: Intuition

- Suppose I told you I had a random number generator that produced a “1” or a “2” with 50% probability each
- You agreed to play a game where you give me \$5 when a 1 appears and I give you \$5 when a 2 appears
- After 30 games, the “random” number generator has produced 29 ones and only 1 two.
- You accuse me of cheating - but what is your evidence?
- You might say something like this: If there was indeed a 50-50 chance, then the probability that we would see 29 ones and only one two is only 0.00001%. Therefore, I reject your claim that the machine is fair.

This is basically exactly what hypothesis testing does - we suppose a claim is true, then evaluate the realized data to test how likely the original claim is

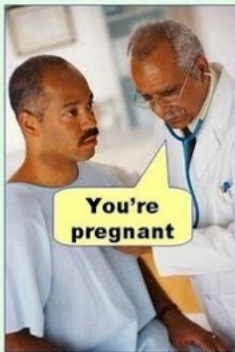
Hypothesis Testing: Type 1 and Type 2 Errors

Choice of a significance level (α) is essentially choosing the trade-off between Type 1 and Type 2 errors

- **Type 1 Error:** This is an error where we reject the null when the null was actually true. In science, this type of error is seen as especially damaging, therefore we try to minimize it. In social science (e.g. economics) a 5% Type 1 error rate has become standard. In disciplines like physics, the bar is much higher (e.g. 0.001% type 1 error rate).
- **Type 2 Error:** This error occurs when we fail to reject the null when the null is actually false. Obviously we want this to be as low as possible, but not at the expense of increasing Type 1 errors beyond the accepted level.

Hypothesis Testing: Type 1 and Type 2 Errors

Type I error
(false positive)

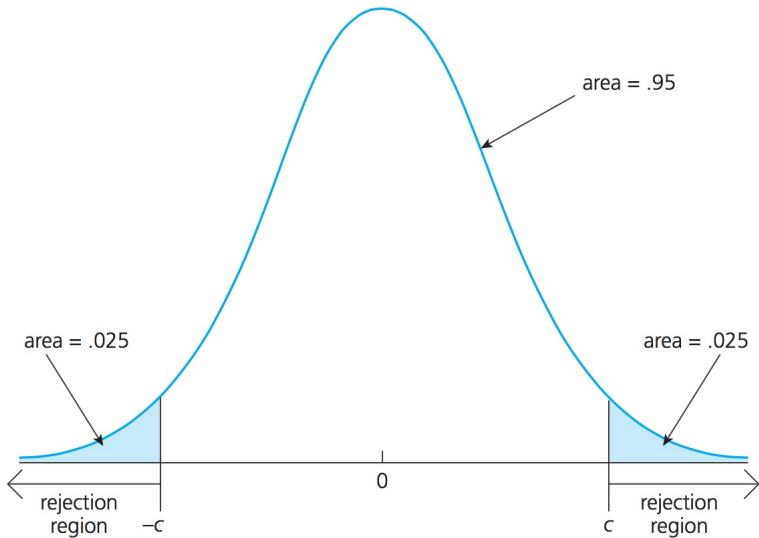


Type II error
(false negative)

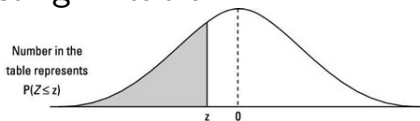


$(H_0 : \text{Not pregnant})$

Rejection region for a 5% significance level test against the two-sided alternative $H_1: \mu \neq \mu_0$.



Hypothesis Testing: z-table



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170

Z-table practice

- 1 $P(Z \geq 1.75) = ?$
- 2 $P(Z > a) = 0.05$, what is a ?
- 3 $P(Z > a) = 0.10$, what is a ?
- 4 $P(Z < -2.32) = ?$
- 5 $P(-a \leq Z \leq a) = 0.90$, what is a ?

Z-table practice

- 1 $P(Z \geq 1.75) = 0.04$
- 2 $P(Z > a) = 0.05$, then $a = 1.64$
- 3 $P(Z > a) = 0.10$, then $a = 1.28$
- 4 $P(Z < -2.32) = 0.01$
- 5 $P(-a \leq Z \leq a) = 0.90$, then $a = 1.64$

Note: That the a for $P(Z > a) = 0.05$ is the same as the a for $P(-a \leq Z \leq a) = 0.90$. Why?

Hypothesis Testing: t-table

TABLE B: t-DISTRIBUTION CRITICAL VALUES

df	Tail probability p											
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460

T-table practice

- 1 $P(t_{50} > c) = 0.05$ what is c ?
- 2 $P(-c > t_{20} \text{ or } t_{20} > c) = 0.05$ what is c ?

T-table practice

- 1 $P(t_{50} > c) = 0.05$ then $c = 1.675$
- 2 $P(-c > t_{20} \text{ or } t_{20} > c) = 0.05$ then $c = 2.086$

The test statistic: One mean test

What do we use for our test statistic? We know from the central limit theorem that

$$\bar{X} \sim N(\mu, \text{Var}(\bar{X}))$$

$$\frac{\bar{X} - \mu}{\text{sd}(\bar{X})} \sim N(0, 1)$$

$$\frac{\bar{X} - \mu}{\text{se}(\bar{X})} \sim t_{n-1}(0, 1)$$

We use this transformation of our estimator \bar{X} as our test statistic:

$$t - \text{stat} = \frac{\bar{X} - \mu}{\text{se}(\bar{X})} = \frac{\bar{X} - \mu}{s_X / \sqrt{n}}$$

The test statistic: One mean test

$$t = \frac{\bar{X} - \mu}{se(\bar{X})} = \frac{\bar{X} - \mu}{s_X / \sqrt{n}}$$

Why is this useful and intuitive?

- The numerator is the deviation between the estimator and our hypothesized population parameter
- We “normalize” this difference based on the standard error of our estimator, so we can think about how “surprising” the estimator would be if the null were true
- This test stat has a known (t) distribution (we call this one a t-stat) so we can easily compare it to critical values

In general, most test statistics have this same basic structure

Hypothesis Testing Process: One Mean Test

Let's say we want to test hypothesis that the true mean of UCB student's GPA is 3.1

① Define hypotheses:

- Since we don't have a good reason to think the average GPA should be higher or lower than 3.1, our hypotheses are:

$$H_0 : \mu = 3.1$$

$$H_1 : \mu \neq 3.1$$

- Note that the \neq in the alternative hypothesis means we are running a **two-sided** test

Hypothesis Testing Process: One Mean Test

- ② Compute the test statistic. Lets assume that we draw a sample of $n = 101$ students, and we calculate the sample estimators:

- $\bar{X} = 2.984$
- $s_x = 0.3723$

The we calculate:

$$t = \frac{\bar{X} - \mu}{\frac{s_x}{\sqrt{n}}} = \frac{2.984 - 3.1}{\frac{0.3723}{\sqrt{101}}} = -3.13$$

Note that we inserted the value of our null hypothesis for μ as “true” population parameter

Hypothesis Testing Process: One Mean Test

- 3 Choose a significance level and find critical value.
 - Let's choose $\alpha = 0.05$
 - Look at t-table for critical value with chosen significance level for a two-tailed test and $n - 1 = 100$ degrees of freedom
 - Make sure I remember whether this is a two sided or one-sided test

TABLE B: t -DISTRIBUTION CRITICAL VALUES

df	Tail probability p											
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
∞	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291

TABLE B: t -DISTRIBUTION CRITICAL VALUES

df	Tail probability p											
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
∞	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.081	3.291

Hypothesis Testing Process Example

- 3 Choose a significance level and find critical value.
 - Let's choose $\alpha = 0.05$
 - Look at t-table for critical value with chosen significance level for a two-tailed test and $n - 1 = 100$ degrees of freedom
 - Make sure I remember whether this is a two sided or one-sided test

$$c = 1.984$$

Hypothesis Testing Process Example

- 4 Do we reject the null? We only reject if:

$$|t| > |c|$$

And we have

$$|-3.13| > |1.984|$$

So we reject the null!

- 5 How do we interpret this:
- We reject the null hypothesis. There is statistical evidence at the 5% level that the average student GPA is different from 3.1
 - What if we had *failed* to reject the null? We would say :
“There is no statistical evidence at the 5% level that average GPA is different from 3.1”

Two-tailed vs. One-tailed

What if we had instead posed the hypotheses

$$H_0 : \mu = 3.1$$

$$H_1 : \mu < 3.1$$

This is a **one-tailed** test. We are only concerned with determining if the GPA is *lower* than 3.1

Let's go back to the t-table from before

- What is the new critical value for our test?

Two-tailed vs. One-tailed

What if we had instead posed the hypotheses

$$H_0 : \mu = 3.1$$

$$H_1 : \mu < 3.1$$

This is a **one-tailed** test. We are only concerned with determining if the GPA is *lower* than 3.1

Let's go back to the t-table from before

- What is the new critical value for our test?
- $c = -1.660$

Helpful rule:

- If two-sided, look for column with value $\alpha/2$
- If one-sided, look for column with value α

Two-tailed vs. One-tailed

What if we had instead posed the hypotheses

$$H_0 : \mu = 3.1$$

$$H_1 : \mu < 3.1$$

- What is the new rejection rule for our test?
- We only reject if:

$$t < c$$

And we have

$$-3.13 < -1.660$$

- Therefore we reject \rightarrow when running a one sided test, the sign of the critical value and the test statistic matters!

Hypothesis Testing: One Mean Test Binary Variables

Let's say we wanted to test a hypothesis on a **binary** variable (e.g. a yes / no vote). Conceptually the process is the same as before, but we have two important (related) differences:

- 1 When we do hypothesis testing, we **assume** the null is true - i.e. that $p = .65$. Now recall the formula for the variance of our estimator \hat{X} when X is binary:

$$\text{Var}(\bar{X}) = \frac{p(1-p)}{n}$$

Notice how this variance **only** depends on the population parameter p .

Because we are assuming that p is equal to our null hypothesis, this implies that we **DON'T** have to estimate the variance!

Hypothesis Testing: One Mean Test Binary Variables

The second difference follows from above:

- ② Because we don't estimate the variance, we never lose the normal distribution for the test statistic! Therefore:

$$z = \frac{\hat{p} - p}{se(\hat{p})} \sim N(0, 1)$$

This implies that we need to look at a z table to find the critical values when dealing with a binary random variable

Hypothesis Testing Binary Variable Example

We want to test that in a yes / no vote 65% of people voted “yes” or if the true proportion is higher than 65%. We sample 200 people and we find that 115 voted yes.

1

$$H_0 : p = 0.65$$

$$H_1 : p > 0.65$$

2 Calculate our test statistic:

- $\hat{p} = \frac{115}{200} = 0.575$
- $Var(\hat{p}) = \frac{p(1-p)}{n} = \frac{.65(1-.65)}{200} = 0.0011375$
- $z = \frac{\hat{p}-p}{se(\hat{p})} = \frac{.575-.65}{0.033727} = -2.224$

Hypothesis Testing Binary Variable Example

- Let's choose a 1% significance level. Using the z-table, for 1% significance and 199 degrees of freedom that

$$c = 2.32$$

- Reject if $z > c$
 - We have $-2.224 < 2.32$, therefore, here we *fail to reject*
- We fail to reject the null hypothesis (at the 1% significance level) that the proportion of people who voted yes is 65%

Hypothesis Testing: Difference in Means

Let's say we wanted to test if two separate groups have different means of some outcome. For example, is income different between rural and urban areas? How does this change our process?

- 1 The null and alternative hypothesis are now

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

We can also rewrite this null and alternative in this way:

$$H_0 : D = \mu_1 - \mu_2 = 0$$

$$H_1 : D = \mu_1 - \mu_2 \neq 0$$

Note: We now are comparing *two* random variables to each other, this complicates things

Hypothesis Testing: Difference in Means

- 2 To calculate our test statistic we know need to calculate \hat{D} as well as find $se(\hat{D})$. \hat{D} is straight forward:

$$\hat{D} = \overline{X_U} - \overline{X_R}$$

In the denominator, we now need to divide by the standard error of $\overline{X_U} - \overline{X_R}$ ($se(\hat{D})$). This is slightly more complicated because *both* of these values are random variables

Hypothesis Testing: Difference in Means

Let's figure out $se(\hat{D})$. We know that

$$\bar{X}_U \sim N\left(\mu_u, \frac{\sigma_u^2}{n_u}\right)$$

$$\bar{X}_R \sim N\left(\mu_r, \frac{\sigma_r^2}{n_r}\right)$$

Then we take the difference:

$$D = \bar{X}_U - \bar{X}_R \sim N\left(\mu_u - \mu_r, \frac{\sigma_u^2}{n_u} + \frac{\sigma_r^2}{n_r}\right)$$

Then formula for this standard error is

$$se(\bar{X}_U - \bar{X}_R) = \sqrt{\frac{s_u^2}{n_u} + \frac{s_r^2}{n_r}}$$

Hypothesis Testing: Difference in Means

Let's say that in our data we calculate

- **Urban:** $\bar{X}_U = 7061.63$, $s_u = 5198$, and $n_u = 1112$
- **Rural:** $\bar{X}_R = 3661.3$, $s_r = 8974$, and $n_r = 1112$

Therefore, the standard error is:

$$se(\hat{D}) = \sqrt{\frac{5198^2}{1112} + \frac{8974^2}{1112}} = 310$$

and

$$\hat{D} = 7061.63 - 3661.3 = 3400.33$$

Hypothesis Testing: Difference in Means

Finally we can calculate the test statistic. Note that since we have to replace σ with the estimator s , we have again moved to a t distribution:

$$t = \frac{\hat{D} - D_{H_0}}{se(\hat{D})} = \frac{3400.33}{310} = 10.97 \sim t_{n_u+n_r-2=1112+1112-2}$$

- **NOTE:** The degrees of freedom for our t-stat (distributed t) is equal to $n_u + n_r - 2$

Hypothesis Testing: Difference in Means

$$t = \frac{\hat{D} - D_{H_0}}{se(\hat{D})} = \frac{3400.33 - 0}{310} = 10.97 \sim t_{n_u+n_r-2=1112+1112-2}$$

- 3 Let's choose significance level of 5%. We can use the normal table because our $DF = 2,222$ is big. The two-tailed critical value is then

$$c = 1.96$$

- 4 Reject if $|t| > |c|$, we have $10.97 > 1.96$, so we reject the null
- 5 Interpret: We reject the null hypothesis. There is statistical evidence that mean incomes are not same in rural and urban areas at the 5% level

Hypothesis Testing: Diff in Means Binary Vars

How does this process change when we have binary variables. Let's say we want to evaluate whether a poverty alleviation program in Mexico increased asset ownership. We see in the data that 14.47% of 14,846 control families own a fridge while 16.53% of 9,213 treatment families do.

We want to test whether the proportion of people who own a refrigerator is higher in the treatment than the control group:

- 1 State our hypotheses

$$H_0 : p_t - p_c = D = 0$$

$$H_1 : p_t - p_c = D > 0$$

Hypothesis Testing: Diff in Means Binary Vars

- 2 Calculate the test statistic. We calculate \hat{D} as before

$$\hat{D} = \hat{p}_t - \hat{p}_c = 0.1653 - 0.1447 = 0.0206$$

But what is $se(\hat{D})$? Remember that

$$Var(\hat{p}) = \frac{\hat{p}(1 - \hat{p})}{n}$$

$$Var(\hat{p}_t - \hat{p}_c) = Var(\hat{p}_t) + Var(\hat{p}_c)$$

Which means that

$$se(\hat{p}_t - \hat{p}_c) = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_t} + \frac{\hat{p}(1 - \hat{p})}{n_c}}$$

Hypothesis Testing: Diff in Means Binary Vars

But there's one really tricky thing here. You'll notice that I've dropped the subscription on the \hat{p} in the standard error formula. *Why?*

$$se(\hat{p}_t - \hat{p}_c) = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n_t} + \frac{\hat{p}(1 - \hat{p})}{n_c}}$$

Our null hypothesis pre-supposes that $p_t = p_c$, so in our standard error formula we need to impose this condition. Therefore the \hat{p} we plug in everywhere has to be the same.

What \hat{p} to plug in then? We plug in the **pooled** \hat{p} :

$$\hat{p}_{pooled} = \frac{n_t \cdot \hat{p}_t}{n_t + n_c} + \frac{n_c \cdot \hat{p}_c}{n_t + n_c}$$

Hypothesis Testing: Diff in Means Binary Vars

In our case

$$\hat{p}_{pooled} = \frac{9213 \cdot 0.1653}{9213 + 14846} + \frac{14846 \cdot 0.1447}{9213 + 14846} = 0.1526$$

Which means

$$se(\hat{D}) = \sqrt{\frac{.1526(1 - .1526)}{9213} + \frac{.1526(1 - .1526)}{14846}} = .00477$$

And our test statistic is (z-stat, i.e. normally distributed):

$$z = \frac{\hat{D} - D_{H_0}}{se(\hat{D})} = \frac{.0206 - 0}{.00477} = 4.32$$

Hypothesis Testing: Diff in Means Binary Vars

- 3 Lets choose 5% significance again. For a one-sided test, the critical value is then $c = 1.65$
- 4 We will reject if $z > c$ and we have $4.32 > 1.65$. Therefore, we reject the null hypothesis
- 5 Interpret: We reject the null hypothesis at the 5% level.
There is statistical evidence that the proportion of households that own a refrigerator in the treatment is higher than that in the control group

Hypothesis Testing β

How does all this apply to Econometrics?

- We want to be able to run a hypothesis test for β .
- In general the testing process for β is the same. However in order to get a known distribution for $\hat{\beta}$
- Recall that assumptions MLR1 to MLR4 imply that $E(\hat{\beta}) = \beta$, while assumption MLR5 stating that $Var(u|x) = \sigma^2$ meant that

$$Var(\hat{\beta}_j) = \frac{\sigma^2}{SST_j(1 - R_j^2)}$$

So, we have the mean and variance of $\hat{\beta}$, but what sort of distribution is it?

Hypothesis Testing β

To find the distribution of $\hat{\beta}$, write it's formula and plug in for y :

$$\begin{aligned}\hat{\beta}_1 &= \frac{\text{cov}(x, y)}{\text{var}(x)} \\ &= \frac{1}{\text{var}(x)} \cdot \text{cov}(x, \beta_0 + \beta_1 x + u) \\ &= \frac{1}{\text{var}(x)} \cdot [\beta_1 \text{var}(x) + \text{cov}(x, u)] \\ &= \beta_1 + \frac{1}{SST_x} \sum_i (x_i - \bar{x}) u_i\end{aligned}$$

We see two things:

- 1 $\hat{\beta}$ is biased if there is any correlation between u and x
- 2 The distribution of the estimator $\hat{\beta}$ depends on the distribution of the true error term u

Hypothesis Testing β

What is the distribution of u ? We don't know: this is a *population* characteristic, so we can't ever know this. So instead, we just assume it:

- **MLR6:** The population error u is normally distributed with mean 0 and variance σ^2 : $u \sim N(0, \sigma^2)$
- If we assume MLR1 - 5 and MLR6, we know that

$$\hat{\beta}_j \sim \text{Normal}[\beta_j, \text{Var}(\hat{\beta}_j)]$$

and therefore

$$\frac{\hat{\beta}_j - \beta_j}{\text{sd}(\hat{\beta}_j)} \sim N(0, 1)$$

Hypothesis Testing β

As before we don't actually know $sd(\hat{\beta})$ because we don't observe σ , so we replace it with $\hat{\sigma}$

As before, this means we lose the normality of the distribution and get a t-distribution:

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \sim t(0, 1)_{n-k-1}$$

Note: n = sample size, k = number of explanatory variables (X_j) in the regression

Hypothesis Testing β

Using this very useful fact

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \sim t(0, 1)_{n-k-1}$$

We can

- 1 Create confidence intervals for $\hat{\beta}$
- 2 Test a hypothesis that our true population parameter β_j is equal to any value

Hypothesis Testing β : Example

We have this population model:

$$\log(\text{rent}) = \beta_0 + \beta_1 \log(\text{pop}) + \beta_2 \log(\text{avginc}) + \beta_3 \text{pctstu} + u$$

We want to test if the percent of students in population has no effect on the rental prices using a sample of 2,000

- 1 Hypotheses: $H_0 : \beta_3 = 0$, $H_1 : \beta_3 \neq 0$.
- 2 To get calculate the test statistic, we need to find the find $\hat{\beta}$ as well as $se(\hat{\beta})$.

Hypothesis Testing β : Example

Both of these are in the Stata output:

lnrent	Coef.	Std. Err.
-----+-----		
lnpop	.0384728	.0432981
lnavginc	.8225076	.0729427
pctstu	.0071755	.0016788
_cons	-2.922981	.7503302

$$t = \frac{0.0071755 - 0}{0.0016788} = 4.27$$

- 3 Significance level and critical value: 5%, and then $c = 1.96$
- 4 We reject the null because $|4.27| > |1.96|$
- 5 At the 5% significance level, we can reject the null that the percentage of students has no effect on predicted rents holding population and income constant.

Hypothesis Testing β : Example II

$$\log(\text{wage}) = \beta_0 + \beta_1 \text{educ} + \beta_k \text{exp} + \dots + u$$

Want to test if the return to education on wages is greater than 5%. Let's say in Stata you see that $\hat{\beta}_{educ} = 0.06$ and $se(\hat{\beta}_{educ}) = 0.0045$ and $n = 50$

- 1 State hypotheses

$$H_0 : \beta_{educ} = 0.05$$

$$H_1 : \beta_{educ} > 0.05$$

Hypothesis Testing β : Example II

- 2 Construct t-stat

$$\frac{\hat{\beta}_{educ} - \beta_{educ}}{se(\hat{\beta}_{exp})} \sim t_{n-k-1}$$
$$\frac{0.06 - 0.05}{0.0045} = 2.22 \sim t_{50-2-1}$$

- 3 Choose 1% significance level. The one sided c value for 1% significance and 47 DF is 2.4
- 4 Reject if $t > c$, and we have $2.22 < 2.4$ therefore we fail to reject
- 5 Interpret: We fail to reject the null that the return to education is 5% at the 1% level