Module 11b – Inference for Two Means (independent samples)

Review Against All Odds: Unit 27 (Inference)

Stat Procedure Diagram – Where we are

		Descriptive Statistics (Describing Pops or Samples)		Inferential Statistics (from Samples)
	Variable Types	Display	Describe	Estimation
Univariate	categorical (nominal or ordinal)*	Bar Graph/Pie Chart	Counts/Percentages	*When binary/dichotomous:* Confidence interval for proportions
	quantitative/continuous	Histogram/Stem & Leaf Box Plot	Mean/St Dev (normal) Median/Min, Q1, Q3, Max (skewed)	Confidence interval for means
		Display	Describe	Significance Tests/Hypothesis Tests
Bivariate	2 categorical	Tables or Bar Graphs	Two-way tables/Crosstabulation	Chi-square test (for goodness of fit)
	1 categorical, 1 quant.	Bar Graphs	Comparison of means/averages	T-test (one sample/group, two samples/groups) ANOVA (two or more samples/groups)
	2 quant.	Scatterplot	Correlation Coef. (Coef. of determ)/ Regression Line	T-test for correlation
		Display	Describe	Significance Tests/Hypothesis Tests
Multivariate	Response Variable is Quant.	-	Ordinary Least Squares Regression (OLS)	F-test for overall model T-tests for each explanatory variable
	Response Variable is categorical (dichotomous)	-	Logistic Regression	Chi-square tests of significance

NOTE: Items highlighted in yellow are covered in this course.

**When a categorical variable has two categories, it is called dichotomous.*

Still focused on…
…*quantitative/continuous variables (inference)*

- Mean Comparisons (with ***unknown*** population standard deviation)

- Types (all are categorical IV and quantitative/continuous DV)

  - One sample/group comparison
  - Two sample/group comparison (matched pairs/before & after)
  - **Two sample/group comparison (independent samples)**
  - Two or more sample/group comparison (ANOVA)

Two sample t-test (independent samples)

- There are two techniques…
  - 1. Assuming equal variances (use pooled variance)
  - 2. Assuming unequal variances (use separate variance)

- We run a test for equality of variances to determine which technique is appropriate.

- *For our purposes (consistent with the text) we will assume unequal variances in our problems.*

- Since we assume unequal variances…
  - Our degrees of freedom (df) will be the **smaller of n_1-1 and n_2-1** (where n_1 is the size of one sample and n_2 is the size of the other sample)

Two sample t-test (independent samples)

- Formula for independent samples t-test (with unequal variances)

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}}}$$

- The numerator is the difference in sample means.
- The denominator is our standard error for the two samples.
- Our t tells us the number of standard errors that separate our two sample average.
- If this value is greater than our critical value (at alpha=.05) we can reject the null (we have good evidence that the two population averages are different)

Two sample t-test (independent samples practice)

- Tips on good weather days vs bad weather days
- A random ***sample*** of 20 receipts from "good weather" days (Sample 1)
- A random ***sample*** of 20 receipts from "bad weather" days (Sample 2)
- [$\bar{x}_1 = 22.22\%$ $\bar{x}_2 = 18.19\%$] [$s_1 = 1.95\%$ $s_2: 2.10\%$]

- **Null Hyp. H_0**: The average for the good and bad weather days are equal
- **Alt Hyp. H_a:** Avg (good weather day) > avg (bad weather day) (1-sided hyp)

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}}}$$

Two sample t-test (independent samples practice)

Software calculates our t-statistic:

t-Test: Two-Sample Assuming Unequal Variances		
	Variable 1	Variable 2
Mean	22.2	18.2
Variance	3.8	4.4
Observations	20	20
t Stat	6.27	

We use degrees of freedom (n-1) for the smaller sample size: (20-1=19)

To determine statistical significance, we find our critical value for df=19 and alpha level .05 for a one-sided test.

Two sample t-test (matched pairs practice)

5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	0.678	0.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	0.677	0.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	0.675	0.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
z^*	0.674	0.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
One-sided P	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
Two-sided P	.50	.40	.30	.20	.10	.05	.04	.02	.01	.005	.002	.001

Two sample t-test (independent samples practice)

- t = 6.27
  - The critical value for a one-sided hypothesis at alpha .05 and df=19 is 1.729
  - Because 6.27 > 1.729 we know that the p-value is less than .05, and therefore ***we reject the null.***

Significance test results

for t-test if alpha=.05 and df = 19