Module 6

General Sir John Kotelawala Defense University

Department of Mathematics

MA 3102 APPLIED STATISTICS

Hypothesis Tests

Introduction

Hypothesis test for population mean one sample

Hypothesis test for difference of population means

Hypothesis test for proportions

Hypothesis test for paired data

The F test for equality of variance

Chi-square test

Hypothesis test for difference of population means

A researcher is interested in comparing the growth response of rice plants against two different levels of the hormone of indoleacetic acid (IAA). He used 2 different hormone concentrations (X, Y) to observe the effect of the hormone on the growth of the rice pant. The researcher would like to test whether there is a difference in growth rate between hormones X and Y. He observed following heights after 1 week time of planting.

X in mm	0.8	1.8	1.0	0.1	0.9	1.7	1.0	1.4	0.9	1.2	0.5		
Y in mm	1.0	0.8	1.6	2.6	1.3	1.1	24	1.8	2.5	1.4	1.9	2.0	**1.2**

Hypotheses:

H_0: $\mu_X = \mu_Y$
H_1: $\mu_X \neq \mu_Y$

- **Let $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$ be, respectively, two independent random samples of sizes n and m from the two normal distributions $N(\mu_X, \sigma_X^2)$ and $N(\mu_Y, \sigma_Y^2)$ respectively. If the sample sizes are not large (Say <30), this problem can be difficult.**
- **If we can assume common but unknown variances (say $\sigma^2 = \sigma_X^2 = \sigma_Y^2$)**

Test Statistics:

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{S_p \sqrt{\left[\frac{1}{n} + \frac{1}{m}\right]}}$$

The pooled estimator of the common standard deviation S_p.

$$S_p = \sqrt{\left[\frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}\right]}$$

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\left[\frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}\right]\left[\frac{1}{n} + \frac{1}{m}\right]}}$$

has t distribution with $n+m-2$ degrees of freedom.

The hypothesis H_o will be rejected in favor to H_1 if the observed $|T|$ is greater than $t_{\alpha/2\,(n+m-2)}$.

X in mm	0.8	1.8	1.0	0.1	0.9	1.7	1.0	1.4	0.9	1.2	0.5		
Y in mm	1.0	0.8	1.6	2.6	1.3	1.1	24	1.8	2.5	1.4	1.9	2.0	**1.2**

Hypotheses:

H_0: $\mu_X = \mu_Y$
H_1: $\mu_X \neq \mu_Y$

Test Statistics:

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\left[\frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}\right]\left[\frac{1}{n} + \frac{1}{m}\right]}}$$

$$T = \frac{(1.03 - 1.66) - (\mu_X - \mu_Y)}{\sqrt{\left[\frac{(10)0.24 + (12)35}{11 + 13 - 2}\right]\left[\frac{1}{11} + \frac{1}{13}\right]}} = -2.81$$

has t distribution with n+m-2 *[11+13-2=22]* degrees of freedom.

TABLE A-3 t Distribution: Critical t Values

	Area in One Tail				
	0.005	0.01	0.025	0.05	0.10
	Area in Two Tails				
Degrees of Freedom	0.01	0.02	0.05	0.10	0.20
1	63.657	31.821	12.706	6.314	3.078
2	9.925	6.965	4.303	2.920	1.886
3	5.841	4.541	3.182	2.353	1.638
4	4.604	3.747	2.776	2.132	1.533
5	4.032	3.365	2.571	2.015	1.476
6	3.707	3.143	2.447	1.943	1.440
7	3.499	2.998	2.365	1.895	1.415
8	3.355	2.896	2.306	1.860	1.397
9	3.250	2.821	2.262	1.833	1.383
10	3.169	2.764	2.228	1.812	1.372
11	3.106	2.718	2.201	1.796	1.363
12	3.055	2.681	2.179	1.782	1.356
13	3.012	2.650	2.160	1.771	1.350
14	2.977	2.624	2.145	1.761	1.345
15	2.947	2.602	2.131	1.753	1.341
16	2.921	2.583	2.120	1.746	1.337
17	2.898	2.567	2.110	1.740	1.333
18	2.878	2.552	2.101	1.734	1.330
19	2.861	2.539	2.093	1.729	1.328
20	2.845	2.528	2.086	1.725	1.325
21	2.831	2.518	2.080	1.721	1.323
22	2.819	2.508	2.074	1.717	1.321
23	2.807	2.500	2.069	1.714	1.319
24	2.797	2.492	2.064	1.711	1.318
25	2.787	2.485	2.060	1.708	1.316

Critical Value:

$$t_{\alpha/2\,(n+m-2)} = t_{0.025\,(11+13-2)} = 2.074$$

Decision:

The hypothesis H_o is be rejected in favor to H_1 because the observed $T = 2.81 > 2.074$.

$$\bar{X} = 1.03$$
$$\bar{Y} = 1.66$$

Hormone Y higher growth rate than hormone X

Example 2:

A product is packaged by a machine with 24 filler heads numbered 1 to 24, with the odd-numbered heads on one side of the machine and the even on the other side. Let X and Y equal the fill weights in grams when a package is filled by an odd-numbered head and an even-numbered head, respectively. Assume that the distributions of X and Y are $N(\mu_X, \sigma^2)$ and $N(\mu_Y, \sigma^2)$, respectively, and that X and Y are independent. We would like to test the null hypothesis H_0: $\mu_X - \mu_Y = 0$ against the alternative hypothesis H_1: $\mu_X - \mu_Y \neq 0$. To perform the test, after the machine has been set up and is running, we shall select one package at random from each filler head and weigh it. The test statistic is that given by Equation 8.2-1 with $n = m = 12$. At an $\alpha = 0.10$ significance level, the critical region is $|t| \geq t_{0.05}(22) = 1.717$.

For the $n = 12$ observations of X, namely,

1071	1076	1070	1083	1082	1067
1078	1080	1075	1084	1075	1080

$\overline{x} = 1076.75$ and $s_x^2 = 29.30$. For the $m = 12$ observations of Y, namely,

1074	1069	1075	1067	1068	1079
1082	1064	1070	1073	1072	1075

$\overline{y} = 1072.33$ and $s_y^2 = 26.24$. The calculated value of the test statistic is

$$t = \frac{1076.75 - 1072.33}{\sqrt{\dfrac{11(29.30) + 11(26.24)}{22}\left(\dfrac{1}{12} + \dfrac{1}{12}\right)}} = 2.05.$$

Since

$$|t| = |2.05| = 2.05 > 1.717,$$

the null hypothesis is rejected at an $\alpha = 0.10$ significance level. Note, however, that

$$|t| = 2.05 < 2.074 = t_{0.025}(22),$$

so that the null hypothesis would not be rejected at an $\alpha = 0.05$ significance level. That is, the p-value is between 0.05 and 0.10.

Note:

Hypotheses:

$$H_0\colon \mu_X = \mu_Y$$
$$H_1\colon \mu_X - \mu_Y < 0$$

Test Statistics:

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$$

When σ_X^2 and σ_Y^2 are known.

If the σ_X^2 and σ_Y^2 unknown but the sample sizes are large σ_X^2 and σ_Y^2 can be replaced by S_X^2 and S_Y^2

Example 3

The target thickness for Fruit Flavored Gum and for Fruit Flavored Bubble Gum is 6.7 hundredths of an inch. Let the independent random variables X and Y equal the respective thicknesses of these gums in hundredths of an inch, and assume that their distributions are $N(\mu_X, \sigma_X^2)$ and $N(\mu_Y, \sigma_Y^2)$, respectively. Because bubble gum has more elasticity than regular gum, it seems as if it would be harder to roll it out to the correct thickness. Thus, we shall test the null hypothesis H_0: $\mu_X = \mu_Y$ against the alternative hypothesis H_1: $\mu_X < \mu_Y$, using samples of sizes $n = 50$ and $m = 40$.

Because the variances are unknown and the sample sizes are large, the test statistic that is used is

$$Z = \frac{\overline{X} - \overline{Y}}{\sqrt{\dfrac{S_X^2}{50} + \dfrac{S_Y^2}{40}}}.$$

At an approximate significance level of $\alpha = 0.01$, the critical region is

$$z \le -z_{0.01} = -2.326.$$

The observed values of X were

6.85	6.60	6.70	6.75	6.75	6.90	6.85	6.90	6.70	6.85
6.60	6.70	6.75	6.70	6.70	6.70	6.55	6.60	6.95	6.95
6.80	6.80	6.70	6.75	6.60	6.70	6.65	6.55	6.55	6.60
6.60	6.70	6.80	6.75	6.60	6.75	6.50	6.75	6.70	6.65
6.70	6.70	6.55	6.65	6.60	6.65	6.60	6.65	6.80	6.60

for which $\overline{x} = 6.701$ and $s_x = 0.108$. The observed values of Y were

7.10	7.05	6.70	6.75	6.90	6.90	6.65	6.60	6.55	6.55
6.85	6.90	6.60	6.85	6.95	7.10	6.95	6.90	7.15	7.05
6.70	6.90	6.85	6.95	7.05	6.75	6.90	6.80	6.70	6.75
6.90	6.90	6.70	6.70	6.90	6.90	6.70	6.70	6.90	6.95

for which $\overline{y} = 6.841$ and $s_y = 0.155$. Since the calculated value of the test statistic is

$$z = \frac{6.701 - 6.841}{\sqrt{0.108^2/50 + 0.155^2/40}} = -4.848 < -2.326,$$

the null hypothesis is clearly rejected.

8.2-5. Some nurses in county public health conducted a survey of women who had received inadequate prenatal care. They used information from birth certificates to select mothers for the survey. The mothers selected were divided into two groups: 14 mothers who said they had five or fewer prenatal visits and 14 mothers who said they had six or more prenatal visits. Let X and Y equal the respective birth weights of the babies from these two sets of mothers, and assume that the distribution of X is $N(\mu_X, \sigma^2)$ and the distribution of Y is $N(\mu_Y, \sigma^2)$.

(a) Define the test statistic and critical region for testing H_0: $\mu_X - \mu_Y = 0$ against H_1: $\mu_X - \mu_Y < 0$. Let $\alpha = 0.05$.

(b) Given that the observations of X were

49	108	110	82	93	114	134
114	96	52	101	114	120	116

and the observations of Y were

133	108	93	119	119	98	106
131	87	153	116	129	97	110

calculate the value of the test statistic and state your conclusion.

Hypothesis test for Proportions

Example

A company produces shelled nuts for sale. Machinery is set so that no more than 10% of nuts are damaged (and profit margins are based on this figure). To test if the machinery is working, a sample of 200 nuts is examined; 29 are damaged. Is the proportion of damaged nuts too high, at the 5% level?

- Here we test $H_0 : p = 0.10$ (the machine is working properly), against $H_A : p > 0.10$ (the machine is not working properly).
- Question? Is $\hat{p} = 29/200 = 0.145$ unusual?
- $Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 \hat{q}_0}{n}}} = \frac{0.145 - 0.10}{\sqrt{\frac{0.1 \times 0.9}{200}}} = \frac{0.045}{0.021213203} = 2.1213$

Example

A company produces shelled nuts for sale. Machinery is set so that no more than 10% of nuts are damaged (and profit margins are based on this figure). To test if the machinery is working, a sample of 200 nuts is examined; 29 are damaged. Is the proportion of damaged nuts too high, at the 5% level?

- Here we test $H_0 : p = 0.10$ (the machine is working properly), against $H_A : p > 0.10$ (the machine is not working properly).
- Question? Is $\hat{p} = 29/200 = 0.145$ unusual?
- $Z = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0 \hat{q}_0}{n}}} = \frac{0.145-0.10}{\sqrt{\frac{0.1\times 0.9}{200}}} = \frac{0.045}{0.021213203} = 2.1213$

Example

Time magazine reported the result of a telephone poll of 800 adult Americans. The question posed of the Americans who were surveyed was: "Should the federal tax on cigarettes be raised to pay for health care reform?" The results of the survey were:

Non-Smokers	Smokers
$n_1 = 605$	$n_2 = 195$
$y_1 = 351$ (yes)	$y_2 = 41$ (yes)
$\hat{p}_1 = \frac{351}{605} = 0.58$	$\hat{p}_2 = \frac{41}{195} = 0.21$

Is there sufficient evidence at the $\alpha = 0.05$ level, say, to conclude that the two populations differ significantly with respect to their opinions?

To test $H_0: \; p_1 = p_2$ against $H_a: \; p_1 \neq p_2$, then calculate the test statistic

- find $\hat{P} = \frac{351+41}{605+195} = \frac{392}{800} = 0.49$ and
- $SE = \sqrt{\hat{p}_1 \hat{q}_1 (1/n_1 + 1/n_2)} = \sqrt{0.49 \times 0.51 (1/605 + 1/195)} = \sqrt{0.001695} = 0.041165$
- Then $Z_0 = \frac{(\hat{p}_1 - \hat{p}_2) - (0)}{SE} = \frac{0.58 - 0.21}{0.041165} = 8.988$

From calculations, $z_0 = 8.988$

- The test is two-sided, $H_a: \; p_1 \neq p_2$.
- At the 5% significance $z_{0.05/2} = 1.96$ is the critical value (2nd last row of p.584 TB).
- Since $z_0 = 8.988 > z_{0.01} = 1.96$, the test is significant.
- At the 5% significance level there is significant sample evidence to reject H_0.
- Decision: Reject H_0 at the 5% significance level.
- **Remark**: For $z_0 = 8.988$, the (two-sided) P-value is $1 - 0.9999999999 \approx 0$ which is much smaller than 5%, leading to the same conclusion.

Table E *(continued)*

Cumulative Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952

Gives the cumulative area from the left up to a vertical line above a specific value of *z*.