STAT FOR BIM LECTURE NOTES | HYPOTHESIS TESTING

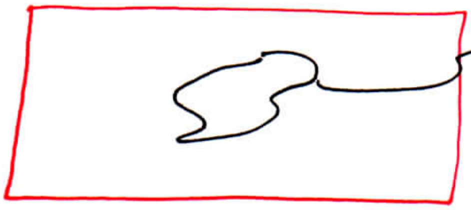
* Hypothesis Testing Questions are YES/NO questions

"Can we conclude that ...?"

"Is there sufficient evidence that ...?"

"Is the claim true?" ... etc.

* We make inference about population parameters.



Random Sample
$n=50$

Population $N=12000$

i.e. Bilkent University Students.

Ex: X: Weekly food expenditure of a student.
y: $=1$ if student makes sport, 0 otherwise.

	Population Parameters (Unknown Constants)	Sample statistics (Known Variables)
MEAN	μ	$\bar{X} = \frac{\sum X_i}{n}$
VARIANCE	σ^2	$s^2 = \frac{\sum X_i^2 - \frac{(\sum X_i)^2}{n}}{n-1}$
PROPORTION	p	$\hat{p} = \frac{\sum y_i}{n}$

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

Hypothesis Testing Steps:

(i) Ho, H_A and α → State what to test

(ii) Test Statistics → Which Table, what is the formula?
Test: One Sample - Two Sample / Mean - Variance - Proportion

(iii) Decision Criteria → When to "Reject Ho"

(iv) Calculation → Calculate Test Statistics

(v) Decision & Conclusion → "Reject Ho" / "Do NOT Reject Ho"

Example// I want to open a Restaurant at Bilkent. I think that opening the restaurant will be profitable if average food expenditure of the students is more than 100TL. Of a random sample of 50 students, mean food expenditure is found to be 103,76 TL. From past experience, variance is known to be 210. At 5% significance level, should I open the restaurant?

Answer// (i) Ho, H_A and α

$H_0: \mu \leq 100$ → Ho and H_A are complementary

$H_A: \mu > 100$ → inequality is always at H_A

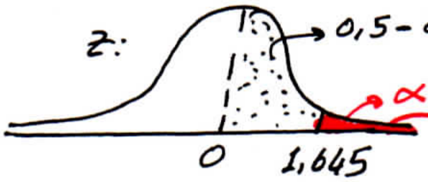
$\alpha = 0,05$ α: Significance Level OR Type-I Error

(ii) Test Statistics

One Sample mean test, σ^2 is known

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

(iii) Decision Criteria



Rejection Region, same side with H_A (we have >)

Reject H_0 if $Z > 1{,}645$

---

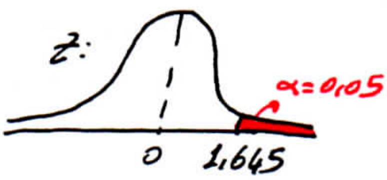
* Rejection Region is the same side with H_A.

Consider the following tests;

$H_0: \mu \leq 100$

$H_A: \mu > 100$

$\alpha = 0{,}05$



Reject H_0 if $Z > 1{,}645$

$H_0: \mu \geq 70$

$H_A: \mu < 70$

$\alpha = 0{,}05$

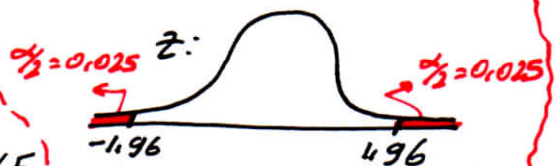
Reject H_0 if $Z < -1{,}645$

(This is called a 2-sided test)

$H_0: \mu = 18$

$H_A: \mu \neq 18$

$\alpha = 0{,}05$



Reject H_0 if

$Z > 1{,}96$ OR $Z < -1{,}96$

(iv) Calculation

$$\left.\begin{array}{l} n = 50 \\ \bar{X} = 103{,}76 \\ \sigma^2 = 210 \end{array}\right\} \Rightarrow Z = \frac{103{,}76 - 100}{\sqrt{210/50}} = 1{,}835$$

(v) Decision & Conclusion

$1{,}835 > 1{,}645$ so we Reject H_0. I can conclude that mean food expenditure is more than 100 TL and open the restaurant at $\alpha = 0{,}05$.

HYPOTHESIS TESTS

One Sample Tests (Parameter vs. NUMBER)

MEAN σ^2 known OR Large n

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

σ^2 unknown AND small n

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} ; \quad df = n-1$$

$$s^2 = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1} ; \quad \bar{x} = \frac{\sum x_i}{n}$$

(For paired sample test, use t and replace $\bar{x}$ with $\bar{D}$; s with s_D)

Two Sample Tests (Parameter-1 vs. Parameter-2)

σ_i^2 known OR Large Samples

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$$

σ_i^2 are unknown, but assumed equal $(\sigma_1^2 = \sigma_2^2)$ AND small samples

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2/n_1 + s_p^2/n_2}}$$

$$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{\underbrace{n_1+n_2-2}_{df}}$$

VARIANCE

$$\chi^2 = \frac{(n-1) \cdot s^2}{\sigma^2}$$

$df = n-1$

Larger Variance

$$F = \frac{s_1^2}{s_2^2}$$

$s_1^2 \rightarrow df = n_1 - 1$

$s_2^2 \rightarrow df = n_2 - 1$

PROPORTION

$$z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \qquad \hat{p} = \frac{x}{n}$$

x: # of success

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{\hat{p}_1 n_1 + \hat{p}_2 n_2}{n_1 + n_2}$$

χ^2 - Tests

Goodness of fit

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

$df = k-1$

Independence

$$\chi^2 = \sum \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

$df = (k-1)(r-1)$

11.77 Tests performed with a random sample of 40 diesel engines produced by a large manufacturer showed that they have a mean thermal efficiency of 31.8 percent with a standard deviation of 2.2 percent. Based on this information and with a probability of a Type I error no greater than 0.05, should the person performing the tests accept the null hypothesis $\mu \geq 32.3$ percent or the alternative hypothesis $\mu < 32.3$ percent?

11.77) $n=40$; $\bar{X}=31,8$; $s=2,2$; $\alpha=0,05$

(i) $H_0: \mu \geq 32,3$

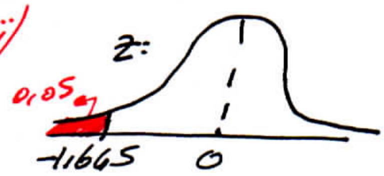
$H_A: \mu < 32,3$

$\alpha = 0,05$

(ii) One Sample Mean test, Large Sample

$$z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$$

(iii) z:



Reject H_0 if $z < -1,645$

(iv) $z = \dfrac{31,8-32,3}{2,2/\sqrt{40}} = -1,437$

(v) Do NOT Reject H_0. Mean efficiency is not less than 32,3 at $\alpha=0,05$

11.82 A random sample of 12 graduates of a secretarial school averaged 73.2 words per minute with a standard deviation of 7.9 words per minute on a typing test. Use the 0.05 level of significance to test the null hypothesis that graduates of this secretarial school average 75.0 words per minute on the given test against the alternative that they average less.

11.82) $n=12$; $\bar{X}=73,2$; $s=7,9$; $\alpha=0,05$

(i) $\mu \geq 75$

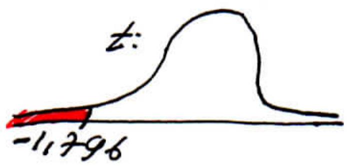
$\mu < 75$

$\alpha = 0,05$

(iii) One sample Mean test, small sample & σ unknown

$$t = \frac{\bar{X}-\mu}{s/\sqrt{n}} ; df = 11$$

(iii) t:



Reject H_0 if $t < -1,796$

(iv) $t = \dfrac{73,2-75}{7,9/\sqrt{12}} = -0,789$

(v) Do NOT Reject H_0.

11.87 An English teacher wants to determine whether the mean reading speed of a certain student is at least 600 words per minute. What can he conclude, if in six 1-minute intervals the student reads 606, 622, 617, 572, 570, and 605 words, and the probability of a Type I error is to be at most 0.05?

11.87)

X_i	X_i^2
606	606^2
622	622^2
617	617^2
572	572^2
570	570^2
+ 605	605^2
$\Sigma x_i = 3592$	$\Sigma x_i^2 = 2152918$

$\boxed{n=6}$

$\bar{X} = \dfrac{3592}{6} = 598,67$

$s^2 = \dfrac{2152918 - \dfrac{3592^2}{6}}{5}$

$= 501,47$

(i) $H_0 : \mu \geq 600$

$H_A : \mu < 600$

$\alpha = 0,05$

(ii) One Sample Mean test, small sample & σ unknown

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}} \; ; \; df = 5$$

(iii)

Reject Ho if $t < -2,015$

(iv) $t = \dfrac{598,67 - 600}{\sqrt{501,47/6}} = -1,145$

(v) Do NOT Reject Ho.

11.111 The following data were obtained in an experiment designed to check whether there is a systematic difference in the weights (in grams) obtained with two different scales:

Rock specimen	Scale I	Scale II
1	12.13	12.17
2	17.56	17.61
3	9.33	9.35
4	11.40	11.42
5	28.62	28.61
6	10.25	10.27
7	23.37	23.42
8	16.27	16.26
9	12.40	12.45
10	24.78	24.75

Use the 0.01 level of significance to test whether the difference between the means of the weights obtained with the two scales is significant.

$\boxed{n = 10}$

11.111)

Rock Specimen	d_i	d_i^2
1	−0,04	$(-0,04)^2$
2	−0,05	$(-0,05)^2$
3	−0,02	
4	−0,02	|
5	0,01	'
6	−0,02	'
7	−0,05	'
8	0,01	'
9	−0,05	'
10	+ 0,03	$0,03^2$

$\Sigma d_i = -0,20$ $\qquad 0,0114 = \Sigma d_i^2$

$$\bar{D} = \frac{\Sigma d_i}{n} = \frac{-20}{10} = -0,02$$

$$s_d^2 = \frac{\Sigma d_i^2 - (\Sigma d_i)^2/n}{n-1} = \frac{0,0114 - (-0,20)^2/10}{9} = 0,00082$$

$$s_d = \sqrt{0,00082} = 0,0287$$

(i) $H_0 : \mu_d = 0$

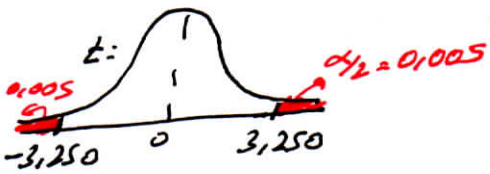
$H_A : \mu_d \neq 0$

$\alpha = 0,01$

(ii) Paired sample test

$$t = \frac{\bar{D} - \mu_d}{s_d/\sqrt{n}} \; ; \; df = 9$$

(iii)



Reject Ho if $|t| > 3,250$

(vi) $t = \dfrac{-0,02 - 0}{0,0287/\sqrt{10}} = -2,204$

(v) Do NOT Reject Ho.

11.103 Twelve measurements each of the hydrogen content (in percent number of atoms) of gases collected from the eruption of two volcanos yielded $\bar{x}_1 = 41.2$, $\bar{x}_2 = 45.8$, $s_1 = 5.2$, and $s_2 = 6.7$. Decide, at the 0.05 level of significance, whether to accept or reject the null hypothesis that there is no difference (with regard to hydrogen content) in the composition of the gases in the two eruptions.

11.103) Volcano I — Volcano II

$\bar{X}_1 = 41,2$ $\quad$ $\bar{X}_2 = 45,8$

$s_1 = 5,2$ $\quad$ $s_2 = 6,7$

$n_1 = 12$ $\quad$ $n_2 = 12$

(i) $H_0: \mu_1 = \mu_2$

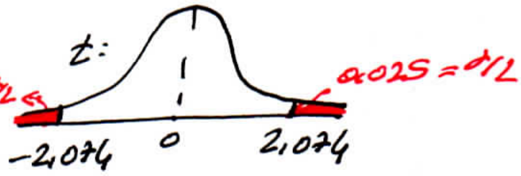
$H_A: \mu_1 \neq \mu_2$

$\alpha = 0,05$

(ii) Two Sample Mean test, small samples, σ_T^2 unknown

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \; ; \; df = 12 + 12 - 2 = 22$$

(iii)



Reject H_0 if $|t| > 2,074$

(iv) $$s_p^2 = \frac{11 \cdot 5,2^2 + 11 \cdot 6,7^2}{22} = 35,45$$

$$t = \frac{41,2 - 45,8}{\sqrt{35,45 \cdot \left(\frac{1}{12} + \frac{1}{12}\right)}} = -1,822$$

(v) Do NOT Reject H_0.

11.122 In a study of the relationship between family size and intelligence, 40 "only children" had an average IQ of 101.5 with standard deviation of 6.7 and 50 "first-borns" in two-child families had an average IQ of 105.9 with a standard deviation of 5.8. Use the 0.05 level of significance to test whether the difference between these two means is significant.

11.122) Only Children — Two Children

$n_1 = 40$ $\quad$ $n_2 = 50$

$\bar{X}_1 = 101,5$ $\quad$ $\bar{X}_2 = 105,9$

$s_1 = 6,7$ $\quad$ $s_2 = 5,8$

(i) $H_0: \mu_1 = \mu_2$

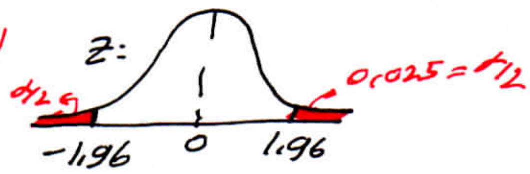
$H_A: \mu_1 \neq \mu_2$

$\alpha = 0,05$

(ii) Two Sample Mean test, Large Samples.

$$z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

(iii)



Reject H_0 if $|z| > 1,96$

(iv) $$z = \frac{101,5 - 105,9}{\sqrt{\frac{6,7^2}{40} + \frac{5,8^2}{50}}} = -3,284$$

(v) Reject H_0. Mean difference is significant at $\alpha = 0,05$

12.18 Past data indicate that the standard deviation of measurements made on sheet metal stampings by experienced inspectors is 0.41 square inch. If a new inspector measures 60 stampings with a standard deviation of 0.48 square inch, test at the 0.05 level of significance whether he is making satisfactory measurements or whether the variability of his measurements is excessive.

12.18) $\sigma = 0{,}41$; $n = 60$; $s = 0{,}48$; $\alpha = 0{,}05$

(i) $H_0 : \sigma^2 \leq 0{,}41$

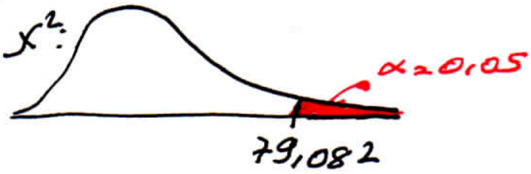
$H_A : \sigma^2 > 0{,}41$

$\alpha = 0{,}05$

(ii) One Sample Variance Test;

$$\chi^2 = \frac{(n-1).s^2}{\sigma^2} ; df = 59$$

(iii)



Reject H_0 if $\chi^2 > 79{,}082$

(iv) $\chi^2 = \dfrac{59 . 0{,}48}{0{,}41} = 69{,}073$

(v) Do NOT Reject H_0.

12.23 Two different lighting techniques are compared by measuring the intensity of light at selected locations in areas lighted by the two methods. If 12 measurements of the first technique have a standard deviation of 2.6 foot-candles and 16 measurements of the second technique have a standard deviation of 4.4 foot-candles, test at the 0.05 level of significance whether the two lighting techniques are equally variable or whether the first technique is less variable than the second.

12.23)

First	Second
$n_1 = 12$	$n_2 = 16$
$s_1 = 2{,}6$	$s_2 = 4{,}4$

(i) $H_0 : \sigma_2^2 \leq \sigma_1^2$

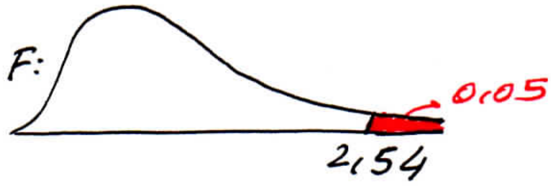
$H_A : \sigma_2^2 > \sigma_1^2$

$\alpha = 0{,}05$

(ii) Two Sample Variance Test

$$F = \frac{\sigma_2^2 \to df = 11}{\sigma_1^2 \to df = 15}$$

(iii)



Reject H_0 if $F > 2{,}54$

(iv) $F = \dfrac{4{,}4^2}{2{,}6^2} = 2{,}864$

(v) Reject H_0.

12.38 In a study of two topical treatments for burn, subjects were scored according to the fraction of rash area that was relieved. The relevant data values are these:

Treatment	Number of subjects	Fraction of area relieved by treatment: Mean	Standard deviation
G	12	0.68	0.21
R	11	0.62	0.09

Using the 10 percent level of significance, test the null hypothesis that the standard deviations of the populations sampled are equal.

12.38) (i) $H_0 : \sigma_1^2 = \sigma_2^2$

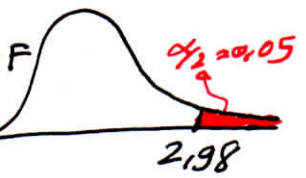
$H_A : \sigma_1^2 \neq \sigma_2^2$

$\alpha = 0{,}10$

(ii) Two Sample Variance Test

$$F = \frac{\sigma_1^2 \to df = 11}{\sigma_2^2 \to df = 10}$$

(iii)



Reject H_0 if $F > 2{,}98$

(iv) $F = \dfrac{0{,}21^2}{0{,}09^2} = 5{,}44$

(v) Reject H_0.

13.42 To check an ambulance service's claim that at least 40 percent of its calls are life-threatening emergencies, a random sample was taken from its files, and it was found that 49 of 150 calls were life-threatening emergencies. If the probability of a Type I error is not to exceed 0.05, what can we conclude about the ambulance service's claim?

(i) $H_0: p \geq 0{,}40$

$H_A: p < 0{,}40$

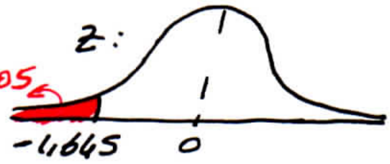
$\alpha = 0{,}05$

(ii) One sample Proportion Test

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

13.42) $\hat{p} = \frac{49}{150} = 0{,}327$; $n = 150$

(iii) z:



Reject H_0 if $z < -1{,}645$

(iv) $$z = \frac{0{,}327 - 0{,}40}{\sqrt{\frac{0{,}40 \cdot 0{,}60}{150}}} = -1{,}825$$

(v) Reject H_0.

13.51 Among 500 marriage license applications, chosen at random twelve years ago, 48 of the women were at least one year older than the men, and among 500 marriage license applications, chosen at random eight years later, 85 of the women were at least one year older than the men. Use the 0.05 level of significance to test whether or not there was an actual increase in the proportion of women on marriage license applications who were at least one year older than the men.

(i) $H_0: p_2 \leq p_1$

$H_A: p_2 > p_1$

$\alpha = 0{,}05$

(ii) Two Sample Proportion Test

$$z = \frac{\hat{p}_2 - \hat{p}_1}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

13.51) 12 years ago | 8 years later

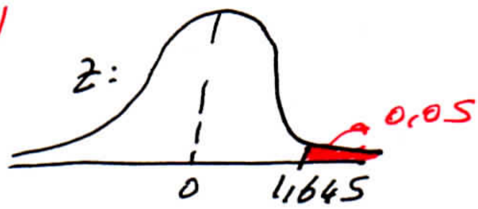
$n_1 = 500$ | $n_2 = 500$

$x_1 = 48$ | $x_2 = 85$

$\hat{p}_1 = \frac{48}{500} = 0{,}096$ | $\hat{p}_2 = \frac{85}{500} = 0{,}170$

$$\hat{p} = \frac{48+85}{500+500} = 0{,}133$$

(iii) z:



Reject H_0 if $z > 1{,}645$

(iv) $$z = \frac{0{,}170 - 0{,}096}{\sqrt{0{,}133 \cdot (1 - 0{,}133)\left(\frac{1}{500} + \frac{1}{500}\right)}}$$

$z = 3{,}646$

(v) Reject H_0.

13.64 A market research organization wants to determine, on the basis of the following information, whether there is a relationship between the size of a tube of toothpaste which a shopper buys and the number of persons in the shopper's household:

		Number of persons			
		1–2	3–4	5–6	7 or more
Size of tube bought	Giant	23	116	78	43
	Large	54	25	16	11
	Small	31	68	39	8

At the 0.01 level of significance, is there a relationship?

13.64)

	1-2	3-4	5-6	7 or more	TOTAL
Giant	23 (54,5)	116 (107,0)	78 (67,1)	43 (31,3)	260
Large	54 (22,9)	28 (44,9)	16 (28,1)	11 (13,1)	109
Small	31 (30,6)	68 (60,1)	39 (37,7)	8 (17,6)	146
TOTAL	108	212	133	62	515

$$E_{ij} = \frac{T_{i.} \times T_{.j}}{T_{..}}$$

$$E_{11} = \frac{260 \cdot 108}{515} = 54,5, \quad E_{12} = \frac{260 \cdot 212}{515} = 107,0$$

- - -

(i) Ho : Size of Tube bought and Family size are INDEPENDENT

HA : There's a relationship between them

$\alpha = 0,01$

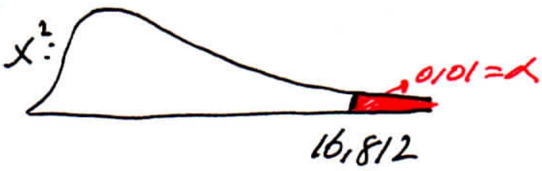
(ii) INDEPENDENCE test

$$\chi^2 = \sum\sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

$df = (3-1).(4-1) = 6$

$(k-1).(c-1)$

(iii) χ^2:



Reject Ho if $\chi^2 > 16,812$

(iv) $$\chi^2 = \frac{(23-54,5)^2}{54,5} + \frac{(116+107,0)^2}{107,0} + \cdots + \frac{(8-17,6)^2}{17,6} = 132,34$$

(v) Reject Ho.

13.81 To see whether a die is balanced, it is rolled 720 times and the following results are obtained: 1 showed 129 times, 2 showed 107 times, 3 showed 98 times, 4 showed 132 times, 5 showed 136 times, and 6 showed 118 times. At the 0.05 level of significance, do these results support the hypothesis that the die is balanced?

13.81) $E_i = p_i \cdot T \quad ; \quad T = 720$

$p_1 = p_2 = \cdots = p_6 = \frac{1}{6}$

$E_i = 720 \cdot \frac{1}{6} = 120$ (if die is balanced)

i	1	2	3	4	5	6
O_i	129	107	98	132	136	118
E_i	120	120	120	120	120	120

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

(i) Die is balanced

(OR $p_1 = p_2 = \dots = p_6 = \frac{1}{6}$)

H_A: Die is NOT balanced

(OR $p_i \neq \frac{1}{6}$ for $\exists\, i = 1, \dots, 6$)

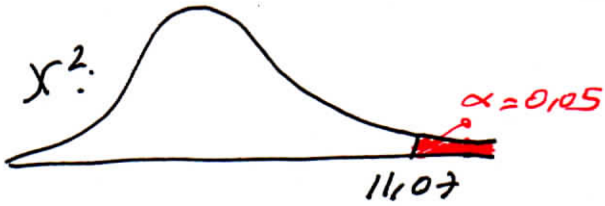
$\alpha = 0{,}05$

(ii) Goodness of Fit test

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

$$df = k - 1 = 6 - 1 = 5$$

(iii) χ^2:



Reject H_0 if $\chi^2 > 11{,}07$

(iv) $$\chi^2 = \frac{(129-120)^2}{120} + \frac{(107-120)^2}{120} + \dots + \frac{(118-120)^2}{120} = 9{,}48$$

(v) Do NOT Reject H_0.

13.86 For 300 consecutive working days, a baker prepares three large chocolate cakes, and those not sold on the same day are donated to a charitable food bank. Given the data shown in the following table, test at the 0.05 level of significance whether they may be looked upon as values of a random variable having a binomial distribution:

Number of cakes sold	*Number of days*
0	2
1	14
2	46
3	238

Use the formula for the binomial distribution to calculate the required probabilities.

13.86) $$\bar{X} = \frac{0.2 + 1.14 + 2.46 + 3.238}{300} = 2{,}73/\text{day}$$

$n\hat{p} = 2{,}73$

$3\hat{p} = 2{,}73 \Rightarrow \hat{p} = 0{,}911$

$X \sim \text{Binomial}(n=3\,;\, p=0{,}911)$

$f(x) = \binom{3}{x} 0{,}911^x \cdot (1-0{,}911)^{3-x}$

$p_0 = \binom{3}{0} 0{,}911^0 \cdot (1-0{,}911)^{3-0} = 0{,}0007$

$p_1 = 0{,}0216\,;\; p_2 = 0{,}2216\,;\; p_3 = 0{,}7561$

i	0	1	2	3	TOTAL
O_i	2	14	46	238	300
E_i	$300 \cdot 0{,}0007 = 0{,}21$	$300 \cdot 0{,}0216 = 6{,}48$	$300 \cdot 0{,}2216 = 66{,}48$	$300 \cdot 0{,}7561 = 226{,}83$	

$$\chi^2 = \frac{(2-0{,}21)^2}{0{,}21} + \frac{(14-6{,}48)^2}{6{,}48} + \frac{(46-66{,}48)^2}{66{,}48} + \frac{(238-226{,}83)^2}{226{,}83} = 30{,}844$$

(i) H_0: Data fits Binomial Distribution.

(OR; $p_0 = 0{,}0007$; $p_1 = 0{,}0216$; $p_2 = 0{,}2216$; $p_3 = 0{,}7501$)

H_A: Data does NOT fit Binomial Distribution

(OR; at least one p_i is different)

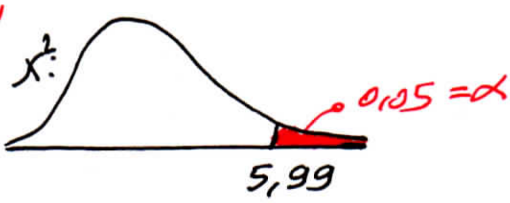
$\alpha = 0{,}05$

(ii) Goodness of Fit test

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} ; \quad df = k - \underbrace{1}_{} - 1 = 4 - 1 - 1 = 2$$

Because we have estimated p from the data

(iii)



Reject H_0 if $\chi^2 > 5{,}99$

(iv) $\chi^2 = 30{,}844$

(v) Reject H_0. Data does NOT fit to Binomial Distribution