Bootcamp – Lesson 3

Hypothesis Testing

Given, *X* is normally distributed then $X \sim N(\mu, \sigma^2)$.

If we take a sample of *n* then the sample mean, $\bar{X}$, is normally distributed,

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$$

For the calculator, we use $\sigma = \frac{\sigma}{\sqrt{n}}$.

We have two methods that can be used:

1. Using p-value to find the probability and then comparing this probability to the significance level.

2. Use the significance level of the test to find the critical value (using inverse normal) and then see if the test statistic (the sample mean) lies inside or outside the critical region.

One-tailed Test

1 The time, X seconds, that it takes Pierre to run a 400 m race can be modelled using $X \sim N(87, 16)$. Pierre changes his diet and claims that the time it takes him to run 400 m has decreased.

(a) Write suitable null and alternative hypotheses to test Pierre's claim.

After changing his diet, Pierre runs 36 separate 400 m races and calculates his mean time on these races to be 86.1 seconds.

(b) Use these 36 races as a sample to test, at the 5% level of significance, whether there is evidence to support Pierre's claim.

2 The mass of a Burmese cat, C, follows a normal distribution with a mean of 4.2 kg and a standard deviation 1.3 kg. Kamala, a cat breeder, claims that Burmese cats weigh more than the average if they live in a household which contains young children. To test her claim, Kamala takes a random sample of 25 cats that live in households containing young children.

The null hypothesis, $H_0: \mu = 4.2$, is used to test Kamala's claim.

(a) (i) Write down the alternative hypothesis to test Kamala's claim.

(ii) Write down the distribution of the sample mean, $\bar{C}$.

(b) Using a 5% level of significance, find the critical region for this test.

Kamala calculates the mean of the 25 cats included in her sample to be 4.65 kg.

(c) Determine the outcome of the hypothesis test at the 5% level of significance, giving your answer in context.

Two-tailed Test

1

The IQ of a student at Calculus High can be modelled as a random variable with the distribution $N(126, 50)$. The headteacher decides to play classical music during lunchtimes and suspects that this has caused a change in the average IQ of the students.

(a) Write suitable null and alternative hypotheses to test the headteacher's suspicion.

The headteacher selects 10 students and asks them to complete an IQ test. Their scores are:

127, 127, 129, 130, 130, 132, 132, 132, 133, 138

(b) Test, at the 5% level of significance, whether there is evidence to support the headteacher's suspicion.

2 A sales representative travels a mean distance of 450 miles each week visiting customers. The company she works for has introduced a new computer system that works out an order in which she visits clients each week and routes to take in order to reduce the mileage. After the first 30 weeks after the introduction of the new system the distance travelled is normally distributed with a mean distance of 435 miles and a standard deviation of 48 miles.

Perform a hypothesis test to test whether the mean distance travelled has changed from the previous figure of 450 miles per week.

Bootcamp – Lesson 3

Hypothesis Testing - PMCC

Correlation Coefficients:

The letter ρ is used to represent the PMCC for a population.

The letter *r* is used to represent the PMCC for a sample.

They both have values between and including -1 and 1.

- ρ or *r* = 0 means no correlation.
- ρ or *r* = 1 means perfect positive correlation.
- ρ or *r* = -1 means perfect negative correlation.

Hypothesis testing:

Testing for positive correlation use $\boldsymbol{H_0}: \rho = 0, \boldsymbol{H_1}: \rho > 0.$

Testing for negative correlation use $\boldsymbol{H_0}: \rho = 0, \boldsymbol{H_1}: \rho < 0.$

Testing for any correlation use $\boldsymbol{H_0}: \rho = 0, \boldsymbol{H_1}: \rho \neq 0.$

We can find the critical value using the statistic tables, page 16.

For the critical value, we need:

1. Significance level
2. One or two tailed
3. n = sample size

One-tailed Test

1 The following table shows the number of hours spent learning to drive, d, and the number of mistakes made in the driving test, m, of ten college students.

d	48	51	51	57	61	68	70	72	73	75
m	19	21	17	12	8	16	7	4	0	1

The product moment correlation coefficient for these data is $r = -0.869$. A driving instructor, Dave, believes there is a negative correlation between the number of hours spent learning to drive and the number of mistakes made in the driving test.

(a) (i) Write down suitable null and alternative hypotheses to test Dave's claim.

(ii) Test, at the 1% level of significance, whether Dave's claim is justified, given

2 Nicole is a Biologist studying the growth of bacteria. She records the number of bacteria on an organism every hour. The table below shows her results for the first eight hours.

Hours (t)	1	2	3	4	5	6	7	8
Number of bacteria (B)	10	50	170	520	1730	5200	17020	58140

Nicole calculates the product moment correlation coefficient as $r = 0.735$.

Nicole claims that there is a positive linear correlation between the number of hours and the number of bacteria.

(a) Test, at the 1% level of significance, whether Nicole's claim is justified. State your hypotheses clearly.

Two-tailed Test

1 Jessica is researching whether there is a correlation between the productivity of university students and the number of hours sleep they get per night.

(a) Write suitable null and alternative hypotheses to test for linear correlation.

Jessica takes a random sample of 25 students, measures their productivity during the day, and records how many hours sleep they had during the previous night. She calculates the product moment correlation coefficient and finds that $r = -0.107$.

Jessica wishes to test, at the 10% level of significance, whether there is evidence that the correlation coefficient for the population is different from zero.

(b) (i) Find the critical regions for Jessica's test.

(ii) Show that, at the 10% level of significance, there is no evidence of a linear correlation.

TABLE 9 CRITICAL VALUES OF THE PRODUCT MOMENT CORRELATION COEFFICIENT

The table gives the critical values, for different significance levels, of the sample product moment correlation coefficient r based on n independent pairs of observations from a bivariate normal distribution with correlation coefficient p = 0.

One tail Two tail n	10% 20%	5% 10%	2.5% 5%	1% 2%	0.5% 1%
4	0.8000	0.9000	0.9500	0.9800	0.9900
5	0.6870	0.8054	0.8783	0.9343	0.9587
6	0.6084	0.7293	0.8114	0.8822	0.9172
7	0.5509	0.6694	0.7545	0.8329	0.8745
8	0.5067	0.6215	0.7067	0.7887	0.8343
9	0.4716	0.5822	0.6664	0.7498	0.7977
10	0.4428	0.5494	0.6319	0.7155	0.7646
11	0.4187	0.5214	0.6021	0.6851	0.7348
12	0.3981	0.4973	0.5760	0.6581	0.7079
13	0.3802	0.4762	0.5529	0.6339	0.6835
14	0.3646	0.4575	0.5324	0.6120	0.6614
15	0.3507	0.4409	0.5140	0.5923	0.6411
16	0.3383	0.4259	0.4973	0.5742	0.6226
17	0.3271	0.4124	0.4821	0.5577	0.6055
18	0.3170	0.4000	0.4683	0.5425	0.5897
19	0.3077	0.3887	0.4555	0.5285	0.5751
20	0.2992	0.3783	0.4438	0.5155	0.5614
21	0.2914	0.3687	0.4329	0.5034	0.5487
22	0.2841	0.3598	0.4227	0.4921	0.5368
23	0.2774	0.3515	0.4132	0.4815	0.5256
24	0.2711	0.3438	0.4044	0.4716	0.5151
25	0.2653	0.3365	0.3961	0.4622	0.5052
26	0.2598	0.3297	0.3882	0.4534	0.4958
27	0.2546	0.3233	0.3809	0.4451	0.4869
28	0.2497	0.3172	0.3739	0.4372	0.4785
29	0.2451	0.3115	0.3673	0.4297	0.4705
30	0.2407	0.3061	0.3610	0.4226	0.4629
31	0.2366	0.3009	0.3550	0.4158	0.4556
32	0.2327	0.2960	0.3494	0.4093	0.4487
33	0.2289	0.2913	0.3440	0.4032	0.4421
34	0.2254	0.2869	0.3388	0.3972	0.4357
35	0.2220	0.2826	0.3338	0.3916	0.4296
36	0.2187	0.2785	0.3291	0.3862	0.4238
37	0.2156	0.2746	0.3246	0.3810	0.4182
38	0.2126	0.2709	0.3202	0.3760	0.4128
39	0.2097	0.2673	0.3160	0.3712	0.4076
40	0.2070	0.2638	0.3120	0.3665	0.4026
41	0.2043	0.2605	0.3081	0.3621	0.3978
42	0.2018	0.2573	0.3044	0.3578	0.3932
43	0.1993	0.2542	0.3008	0.3536	0.3887
44	0.1970	0.2512	0.2973	0.3496	0.3843
45	0.1947	0.2483	0.2940	0.3457	0.3801
46	0.1925	0.2455	0.2907	0.3420	0.3761
47	0.1903	0.2429	0.2876	0.3384	0.3721
48	0.1883	0.2403	0.2845	0.3348	0.3683
49	0.1863	0.2377	0.2816	0.3314	0.3646
50	0.1843	0.2353	0.2787	0.3281	0.3610
60	0.1678	0.2144	0.2542	0.2997	0.3301
70	0.1550	0.1982	0.2352	0.2776	0.3060
80	0.1448	0.1852	0.2199	0.2597	0.2864
90	0.1364	0.1745	0.2072	0.2449	0.2702
100	0.1292	0.1654	0.1966	0.2324	0.2565