Applied Statistics for Life Sciences

Dmitri D. Pervouchine

Centre de Regulació Genòmica

Module 4: Statistical Inference, Part II.

Contents

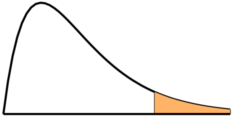
Estimation of Variance

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

$$t = \frac{\overline{X} - \mu}{s/\sqrt{n}}$$

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

Chi-square table



	Tail Probability $P(\chi^2 \geq a)$									
df	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	—	—	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801

Problem 1.1

A supplier of 100 ohm/cm silicon wafers claims that his fabrication process can produce wafers with sufficient consistency so that the standard deviation of resistance for the lot does not exceed 10 ohm/cm. A sample of 10 wafers taken from the lot has a standard deviation of 13.97 ohm/cm. Is the suppliers claim reasonable?

Problem 1.1

A supplier of 100 ohm/cm silicon wafers claims that his fabrication process can produce wafers with sufficient consistency so that the standard deviation of resistance for the lot does not exceed 10 ohm/cm. A sample of 10 wafers taken from the lot has a standard deviation of 13.97 ohm/cm. Is the suppliers claim reasonable?

Solution

- $H_0 : \sigma = 10$

Problem 1.1

A supplier of 100 ohm/cm silicon wafers claims that his fabrication process can produce wafers with sufficient consistency so that the standard deviation of resistance for the lot does not exceed 10 ohm/cm. A sample of 10 wafers taken from the lot has a standard deviation of 13.97 ohm/cm. Is the suppliers claim reasonable?

Solution

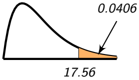
- $H_0 : \sigma = 10$
- $H_a : \sigma > 10$

Problem 1.1

A supplier of 100 ohm/cm silicon wafers claims that his fabrication process can produce wafers with sufficient consistency so that the standard deviation of resistance for the lot does not exceed 10 ohm/cm. A sample of 10 wafers taken from the lot has a standard deviation of 13.97 ohm/cm. Is the suppliers claim reasonable?

Solution

- $H_0 : \sigma = 10$
- $H_a : \sigma > 10$
- $df = 10 - 1 = 9, \quad \mathrm{P}(s^2 > 13.97^2) = \mathrm{P}(\chi^2(9) > \frac{9 * 13.97^2}{10^2}) = \mathrm{P}(\chi^2(9) > 17.56) = 0.0406$

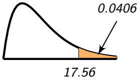


Problem 1.1

A supplier of 100 ohm/cm silicon wafers claims that his fabrication process can produce wafers with sufficient consistency so that the standard deviation of resistance for the lot does not exceed 10 ohm/cm. A sample of 10 wafers taken from the lot has a standard deviation of 13.97 ohm/cm. Is the suppliers claim reasonable?

Solution

- $H_0 : \sigma = 10$
- $H_a : \sigma > 10$
- $df = 10 - 1 = 9, \quad \mathrm{P}(s^2 > 13.97^2) = \mathrm{P}(\chi^2(9) > \frac{9*13.97^2}{10^2}) = \mathrm{P}(\chi^2(9) > 17.56) = 0.0406$



- *At 5% significance level the suppliers claim doesn't seem reasonable, i.e., there is enough reason to believe that* $\sigma > 10$.

Problem 1.2

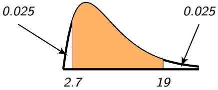
A container of oil is supposed to contain 1000 ml of oil. We want to be sure that the standard deviation of the oil container is less than 20 ml. We randomly select 10 cans of oil with a mean of 997 ml and a standard deviation of 32 ml. Using these sample construct a 95% confidence interval for the true value of sigma. Does the confidence interval suggest that the variation in oil containers is at an acceptable level?

Problem 1.2

A container of oil is supposed to contain 1000 ml of oil. We want to be sure that the standard deviation of the oil container is less than 20 ml. We randomly select 10 cans of oil with a mean of 997 ml and a standard deviation of 32 ml. Using these sample construct a 95% confidence interval for the true value of sigma. Does the confidence interval suggest that the variation in oil containers is at an acceptable level?

Solution

- *Two-sided interval*

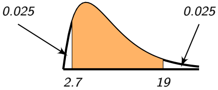


Problem 1.2

A container of oil is supposed to contain 1000 ml of oil. We want to be sure that the standard deviation of the oil container is less than 20 ml. We randomly select 10 cans of oil with a mean of 997 ml and a standard deviation of 32 ml. Using these sample construct a 95% confidence interval for the true value of sigma. Does the confidence interval suggest that the variation in oil containers is at an acceptable level?

Solution

- *Two-sided interval*



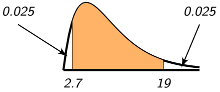
- $P(2.7 < \chi^2(9) < 19) = 0.95$

Problem 1.2

A container of oil is supposed to contain 1000 ml of oil. We want to be sure that the standard deviation of the oil container is less than 20 ml. We randomly select 10 cans of oil with a mean of 997 ml and a standard deviation of 32 ml. Using these sample construct a 95% confidence interval for the true value of sigma. Does the confidence interval suggest that the variation in oil containers is at an acceptable level?

Solution

- *Two-sided interval*



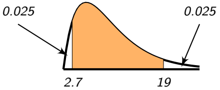
- $P(2.7 < \chi^2(9) < 19) = 0.95$
- $2.7 < \frac{9*32^2}{\sigma^2} < 19$

Problem 1.2

A container of oil is supposed to contain 1000 ml of oil. We want to be sure that the standard deviation of the oil container is less than 20 ml. We randomly select 10 cans of oil with a mean of 997 ml and a standard deviation of 32 ml. Using these sample construct a 95% confidence interval for the true value of sigma. Does the confidence interval suggest that the variation in oil containers is at an acceptable level?

Solution

- *Two-sided interval*



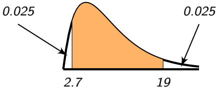
- $P(2.7 < \chi^2(9) < 19) = 0.95$
- $2.7 < \frac{9*32^2}{\sigma^2} < 19$
- $\frac{9*32^2}{19} < \sigma^2 < \frac{9*32^2}{2.7}$

Problem 1.2

A container of oil is supposed to contain 1000 ml of oil. We want to be sure that the standard deviation of the oil container is less than 20 ml. We randomly select 10 cans of oil with a mean of 997 ml and a standard deviation of 32 ml. Using these sample construct a 95% confidence interval for the true value of sigma. Does the confidence interval suggest that the variation in oil containers is at an acceptable level?

Solution

- *Two-sided interval*



- $P(2.7 < \chi^2(9) < 19) = 0.95$
- $2.7 < \frac{9*32^2}{\sigma^2} < 19$
- $\frac{9*32^2}{19} < \sigma^2 < \frac{9*32^2}{2.7}$
- *We are 95% confident that* σ^2 *is between* 22 *and* 58 *ml.*

Fisher F-distribution

- $\frac{s_1^2(n_1-1)}{\sigma_1^2} \sim \chi^2_{n_1-1}$

Fisher F-distribution

- $\frac{s_1^2(n_1-1)}{\sigma_1^2} \sim \chi^2_{n_1-1}$
- $\frac{s_2^2(n_2-1)}{\sigma_2^2} \sim \chi^2_{n_2-1}$

Fisher F-distribution

- $\frac{s_1^2(n_1-1)}{\sigma_1^2} \sim \chi^2_{n_1-1}$
- $\frac{s_2^2(n_2-1)}{\sigma_2^2} \sim \chi^2_{n_2-1}$
- $\frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} \sim \frac{\frac{1}{n_1-1}\chi^2_{n_1-1}}{\frac{1}{n_2-1}\chi^2_{n_2-1}} = F(n_1-1, n_2-1)$

Fisher F-distribution

- $\frac{s_1^2(n_1-1)}{\sigma_1^2} \sim \chi^2_{n_1-1}$
- $\frac{s_2^2(n_2-1)}{\sigma_2^2} \sim \chi^2_{n_2-1}$
- $\frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} \sim \frac{\frac{1}{n_1-1}\chi^2_{n_1-1}}{\frac{1}{n_2-1}\chi^2_{n_2-1}} = F(n_1-1, n_2-1)$
- The F-distribution is the ratio of two independent χ^2 variables divided by their respective degrees of freedom

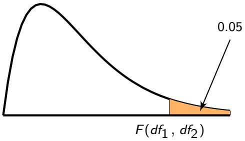
Fisher F-distribution

- $\frac{s_1^2(n_1-1)}{\sigma_1^2} \sim \chi^2_{n_1-1}$
- $\frac{s_2^2(n_2-1)}{\sigma_2^2} \sim \chi^2_{n_2-1}$
- $\frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} \sim \frac{\frac{1}{n_1-1}\chi^2_{n_1-1}}{\frac{1}{n_2-1}\chi^2_{n_2-1}} = F(n_1-1, n_2-1)$
- The F-distribution is the ratio of two independent χ^2 variables divided by their respective degrees of freedom
- The F-test is designed to test if two population variances are equal

$$H_0 : \sigma_1^2 = \sigma_2^2$$

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

Fisher F-distribution



df_1	$df_1 = 2$	2	3	4	5	6	7	8	9	10
1	161.45	18.51	10.13	7.71	6.61	5.99	5.59	5.32	5.12	4.96
2	199.50	19.00	9.55	6.94	5.79	5.14	4.74	4.46	4.26	4.10
3	215.71	19.16	9.28	6.59	5.41	4.76	4.35	4.07	3.86	3.71
4	224.58	19.25	9.12	6.39	5.19	4.53	4.12	3.84	3.63	3.48
5	230.16	19.30	9.01	6.26	5.05	4.39	3.97	3.69	3.48	3.33
6	233.99	19.33	8.94	6.16	4.95	4.28	3.87	3.58	3.37	3.22
7	236.77	19.35	8.89	6.09	4.88	4.21	3.79	3.50	3.29	3.14
8	238.88	19.37	8.85	6.04	4.82	4.15	3.73	3.44	3.23	3.07
9	240.54	19.38	8.81	6.00	4.77	4.10	3.68	3.39	3.18	3.02
10	241.88	19.40	8.79	5.96	4.74	4.06	3.64	3.35	3.14	2.98

$$\mathrm{P}\left(F(df_1, df_2) < x\right) = \mathrm{P}\left(\frac{1}{F(df_1, df_2)} > \frac{1}{x}\right) = \mathrm{P}\left(F(df_2, df_1) > \frac{1}{x}\right)$$

Problem 2.1 (Exercise laboratory problem revisited)

A hospital exercise laboratory technician notes the resting pulse rates of five joggers to be 60, 58, 59, 61, and 67, respectively, while the resting pulse rates of seven non-exercisers are 83, 60, 75, 71, 91, 82, and 84, respectively. The means and standard deviations for these samples are 61, 78, 3.54, and 10.23, respectively. Is equal variances assumption reasonable in this case?

Problem 2.1 (Exercise laboratory problem revisited)

A hospital exercise laboratory technician notes the resting pulse rates of five joggers to be 60, 58, 59, 61, and 67, respectively, while the resting pulse rates of seven non-exercisers are 83, 60, 75, 71, 91, 82, and 84, respectively. The means and standard deviations for these samples are 61, 78, 3.54, and 10.23, respectively. Is equal variances assumption reasonable in this case?

Solution

- $H_0 : \sigma_1^2 = \sigma_2^2$

Problem 2.1 (Exercise laboratory problem revisited)

A hospital exercise laboratory technician notes the resting pulse rates of five joggers to be 60, 58, 59, 61, and 67, respectively, while the resting pulse rates of seven non-exercisers are 83, 60, 75, 71, 91, 82, and 84, respectively. The means and standard deviations for these samples are 61, 78, 3.54, and 10.23, respectively. Is equal variances assumption reasonable in this case?

Solution

- $H_0 : \sigma_1^2 = \sigma_2^2$
- $H_a : \sigma_1^2 \neq \sigma_2^2$

Problem 2.1 (Exercise laboratory problem revisited)

A hospital exercise laboratory technician notes the resting pulse rates of five joggers to be 60, 58, 59, 61, and 67, respectively, while the resting pulse rates of seven non-exercisers are 83, 60, 75, 71, 91, 82, and 84, respectively. The means and standard deviations for these samples are 61, 78, 3.54, and 10.23, respectively. Is equal variances assumption reasonable in this case?

Solution

- $H_0 : \sigma_1^2 = \sigma_2^2$
- $H_a : \sigma_1^2 \neq \sigma_2^2$
- $F = \frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} = \frac{s_1^2}{s_2^2} = \frac{3.54}{10.23} = 0.346$

Problem 2.1 (Exercise laboratory problem revisited)

A hospital exercise laboratory technician notes the resting pulse rates of five joggers to be 60, 58, 59, 61, and 67, respectively, while the resting pulse rates of seven non-exercisers are 83, 60, 75, 71, 91, 82, and 84, respectively. The means and standard deviations for these samples are 61, 78, 3.54, and 10.23, respectively. Is equal variances assumption reasonable in this case?

Solution

- $H_0 : \sigma_1^2 = \sigma_2^2$
- $H_a : \sigma_1^2 \neq \sigma_2^2$
- $F = \frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} = \frac{s_1^2}{s_2^2} = \frac{3.54}{10.23} = 0.346$
- $df_1 = 5 - 1 = 4$; $df_2 = 7 - 1 = 6$

Problem 2.1 (Exercise laboratory problem revisited)

A hospital exercise laboratory technician notes the resting pulse rates of five joggers to be 60, 58, 59, 61, and 67, respectively, while the resting pulse rates of seven non-exercisers are 83, 60, 75, 71, 91, 82, and 84, respectively. The means and standard deviations for these samples are 61, 78, 3.54, and 10.23, respectively. Is equal variances assumption reasonable in this case?

Solution

- $H_0 : \sigma_1^2 = \sigma_2^2$
- $H_a : \sigma_1^2 \neq \sigma_2^2$
- $F = \frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} = \frac{s_1^2}{s_2^2} = \frac{3.54}{10.23} = 0.346$
- $df_1 = 5 - 1 = 4;\ df_2 = 7 - 1 = 6$
- $\mathrm{P}(F(4,6) < 0.346) = \mathrm{P}(F(6,4) > \frac{1}{0.346}) = \mathrm{P}(F(6,4) > 2.89) > 0.05$ *since* $F_{0.05}(6,4) = 6.16$

Problem 2.1 (Exercise laboratory problem revisited)

A hospital exercise laboratory technician notes the resting pulse rates of five joggers to be 60, 58, 59, 61, and 67, respectively, while the resting pulse rates of seven non-exercisers are 83, 60, 75, 71, 91, 82, and 84, respectively. The means and standard deviations for these samples are 61, 78, 3.54, and 10.23, respectively. Is equal variances assumption reasonable in this case?

Solution

- $H_0 : \sigma_1^2 = \sigma_2^2$
- $H_a : \sigma_1^2 \neq \sigma_2^2$
- $F = \frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} = \frac{s_1^2}{s_2^2} = \frac{3.54}{10.23} = 0.346$
- $df_1 = 5 - 1 = 4;\ df_2 = 7 - 1 = 6$
- $\mathrm{P}(F(4,6) < 0.346) = \mathrm{P}(F(6,4) > \frac{1}{0.346}) = \mathrm{P}(F(6,4) > 2.89) > 0.05$ *since* $F_{0.05}(6,4) = 6.16$
- *There is not enough evidence to reject* H_0 *at the 5% significance level, i.e., equal variances assumption is not unreasonable.*

Estimation of Sample Size

- What is a minimum sample size needed to estimate the population mean within 2 units?

- What is a minimum sample size needed to estimate the population proportion within 2 percent units?

Problem 3.1

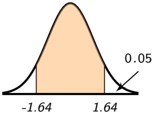
An electrical firm which manufactures a certain type of bulb wants to estimate its mean life. Assuming that the life of the light bulb is normally distributed and that the standard deviation is known to be 40 hours, how many bulbs should be tested so that we can be 90% confident that the estimate of the mean will not differ from the true mean life by more than 10 hours?

Problem 3.1

An electrical firm which manufactures a certain type of bulb wants to estimate its mean life. Assuming that the life of the light bulb is normally distributed and that the standard deviation is known to be 40 hours, how many bulbs should be tested so that we can be 90% confident that the estimate of the mean will not differ from the true mean life by more than 10 hours?

Solution

- $\mu = \overline{X} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$, *where* $z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 10$

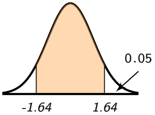


Problem 3.1

An electrical firm which manufactures a certain type of bulb wants to estimate its mean life. Assuming that the life of the light bulb is normally distributed and that the standard deviation is known to be 40 hours, how many bulbs should be tested so that we can be 90% confident that the estimate of the mean will not differ from the true mean life by more than 10 hours?

Solution

- $\mu = \overline{X} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$, *where* $z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 10$



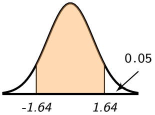
- $1.64 \frac{40}{\sqrt{n}} = 10$

Problem 3.1

An electrical firm which manufactures a certain type of bulb wants to estimate its mean life. Assuming that the life of the light bulb is normally distributed and that the standard deviation is known to be 40 hours, how many bulbs should be tested so that we can be 90% confident that the estimate of the mean will not differ from the true mean life by more than 10 hours?

Solution

- $\mu = \overline{X} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$, *where* $z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 10$



- $1.64 \frac{40}{\sqrt{n}} = 10$
- $n = 43.03 \rightarrow 44$

Problem 3.2

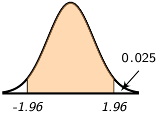
A quality control engineer wants to estimate the fraction of defective bulbs in a large lot of light bulbs. From past experience, he feels that the actual fraction of defective bulbs should be somewhere around 0.2 . How large a sample should be taken if he wants to estimate the true fraction within .02 using a 95% confidence interval?

Problem 3.2

A quality control engineer wants to estimate the fraction of defective bulbs in a large lot of light bulbs. From past experience, he feels that the actual fraction of defective bulbs should be somewhere around 0.2 . How large a sample should be taken if he wants to estimate the true fraction within .02 using a 95% confidence interval?

Solution

- $p = \hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n}}$, *where* $z_{\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n}} = 0.02$

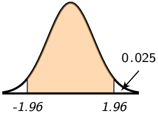


Problem 3.2

A quality control engineer wants to estimate the fraction of defective bulbs in a large lot of light bulbs. From past experience, he feels that the actual fraction of defective bulbs should be somewhere around 0.2 . How large a sample should be taken if he wants to estimate the true fraction within .02 using a 95% confidence interval?

Solution

- $p = \hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n}}$, *where* $z_{\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n}} = 0.02$



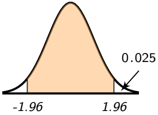
- $1.96\sqrt{\frac{0.2*0.8}{n}} = 0.02$

Problem 3.2

A quality control engineer wants to estimate the fraction of defective bulbs in a large lot of light bulbs. From past experience, he feels that the actual fraction of defective bulbs should be somewhere around 0.2 . How large a sample should be taken if he wants to estimate the true fraction within .02 using a 95% confidence interval?

Solution

- $p = \hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n}}$, *where* $z_{\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n}} = 0.02$



- $1.96\sqrt{\frac{0.2*0.8}{n}} = 0.02$
- $n = 1536.64 \to 1537$

Problem 3.3

Many television viewers express doubts about the validity of certain commercials. Let p represent the true proportion of consumers who believe what is shown in Timex television commercials. If Timex has no prior information regarding the true value of p, how many consumers should be included in their sample so that they will be 85% confident that their estimate is within 0.03 of the true value of p?

Problem 3.3

Many television viewers express doubts about the validity of certain commercials. Let p represent the true proportion of consumers who believe what is shown in Timex television commercials. If Timex has no prior information regarding the true value of p, how many consumers should be included in their sample so that they will be 85% confident that their estimate is within 0.03 of the true value of p?

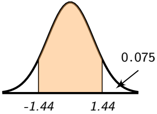
Solution

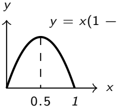
- $p = \hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n}}$, *where* $z_{\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n}} = 0.03$

Problem 3.3

Many television viewers express doubts about the validity of certain commercials. Let p represent the true proportion of consumers who believe what is shown in Timex television commercials. If Timex has no prior information regarding the true value of p, how many consumers should be included in their sample so that they will be 85% confident that their estimate is within 0.03 of the true value of p?

Solution

- $p = \hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n}}$, *where* $z_{\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n}} = 0.03$
- 

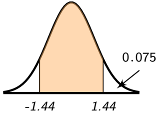


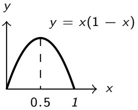
$p = \frac{1}{2}$ *is the "worst" case*

Problem 3.3

Many television viewers express doubts about the validity of certain commercials. Let p represent the true proportion of consumers who believe what is shown in Timex television commercials. If Timex has no prior information regarding the true value of p, how many consumers should be included in their sample so that they will be 85% confident that their estimate is within 0.03 of the true value of p?

Solution

- $p = \hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n}}$, *where* $z_{\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n}} = 0.03$
- 

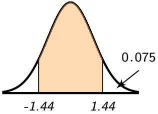


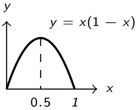
$p = \frac{1}{2}$ *is the "worst" case*
- $1.44\sqrt{\frac{0.5*0.5}{n}} = 0.03$

Problem 3.3

Many television viewers express doubts about the validity of certain commercials. Let p represent the true proportion of consumers who believe what is shown in Timex television commercials. If Timex has no prior information regarding the true value of p, how many consumers should be included in their sample so that they will be 85% confident that their estimate is within 0.03 of the true value of p?

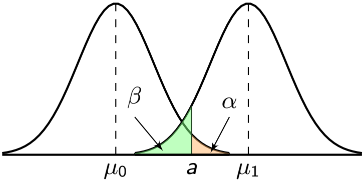
Solution

- $p = \hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n}}$, *where* $z_{\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n}} = 0.03$
- 



$p = \frac{1}{2}$ *is the "worst" case*
- $1.44\sqrt{\frac{0.5*0.5}{n}} = 0.03$
- $n = 576$

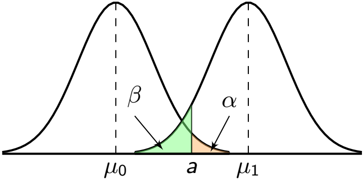
Contribution of type I and type II errors



What is n such that the probability of committing type I error is α and the probability of committing type II error is β? The size of the effect is $\mu_1 - \mu_0 = \Delta$.

- $P(\overline{X} > a|\mu = \mu_0) = \alpha \quad P(\overline{X} < a|\mu = \mu_1) = \beta$

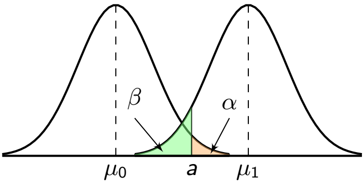
Contribution of type I and type II errors



What is n such that the probability of committing type I error is α and the probability of committing type II error is β? The size of the effect is $\mu_1 - \mu_0 = \Delta$.

- $P(\overline{X} > a | \mu = \mu_0) = \alpha \quad P(\overline{X} < a | \mu = \mu_1) = \beta$
- $\begin{cases} \frac{a - \mu_0}{\sigma/\sqrt{n}} & = z_\alpha \\ \frac{\mu_1 - a}{\sigma/\sqrt{n}} & = z_\beta \end{cases}$

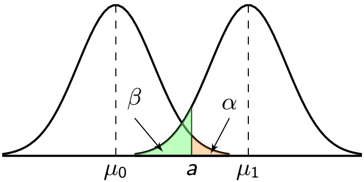
Contribution of type I and type II errors



What is n such that the probability of committing type I error is α and the probability of committing type II error is β? The size of the effect is $\mu_1 - \mu_0 = \Delta$.

- $\mathrm{P}(\overline{X} > a|\mu = \mu_0) = \alpha \quad \mathrm{P}(\overline{X} < a|\mu = \mu_1) = \beta$
- $\begin{cases} \frac{a-\mu_0}{\sigma/\sqrt{n}} & = z_\alpha \\ \frac{\mu_1-a}{\sigma/\sqrt{n}} & = z_\beta \end{cases}$
- $a = \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}} = \mu_1 - z_\beta \frac{\sigma}{\sqrt{n}} \qquad (z_\alpha + z_\beta)\frac{\sigma}{\sqrt{n}} = \mu_1 - \mu_0 = \Delta$

Contribution of type I and type II errors



What is n such that the probability of committing type I error is α and the probability of committing type II error is β? The size of the effect is $\mu_1 - \mu_0 = \Delta$.

- $P(\overline{X} > a|\mu = \mu_0) = \alpha \quad P(\overline{X} < a|\mu = \mu_1) = \beta$
- $\begin{cases} \frac{a-\mu_0}{\sigma/\sqrt{n}} & = z_\alpha \\ \frac{\mu_1 - a}{\sigma/\sqrt{n}} & = z_\beta \end{cases}$
- $a = \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}} = \mu_1 - z_\beta \frac{\sigma}{\sqrt{n}} \qquad (z_\alpha + z_\beta)\frac{\sigma}{\sqrt{n}} = \mu_1 - \mu_0 = \Delta$
- $n = \left(\frac{(z_\alpha + z_\beta)\sigma}{\Delta}\right)^2$

Problem 3.4

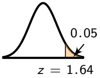
A clinical research organization is to design a pre-clinical of efficacy of a new drug to reduce the cholesterol level. The drug will be commercialized if the reduction of cholesterol be at least 2 mg/dL. Assuming the standard deviation of the cholesterol level in the target population is 20 mg/dL, what is the minimum sample size to achieve the desired reduction with at 5% significance level and with 15% type II error rate (85% power)?

Problem 3.4

A clinical research organization is to design a pre-clinical of efficacy of a new drug to reduce the cholesterol level. The drug will be commercialized if the reduction of cholesterol be at least 2 mg/dL. Assuming the standard deviation of the cholesterol level in the target population is 20 mg/dL, what is the minimum sample size to achieve the desired reduction with at 5% significance level and with 15% type II error rate (85% power)?

Solution

- $z_\alpha = z_{0.05} = 1.64$

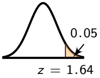


Problem 3.4

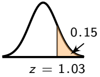
A clinical research organization is to design a pre-clinical of efficacy of a new drug to reduce the cholesterol level. The drug will be commercialized if the reduction of cholesterol be at least 2 mg/dL. Assuming the standard deviation of the cholesterol level in the target population is 20 mg/dL, what is the minimum sample size to achieve the desired reduction with at 5% significance level and with 15% type II error rate (85% power)?

Solution

- $z_\alpha = z_{0.05} = 1.64$



- $z_\beta = z_{0.15} = 1.03$

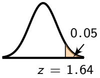


Problem 3.4

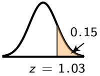
A clinical research organization is to design a pre-clinical of efficacy of a new drug to reduce the cholesterol level. The drug will be commercialized if the reduction of cholesterol be at least 2 mg/dL. Assuming the standard deviation of the cholesterol level in the target population is 20 mg/dL, what is the minimum sample size to achieve the desired reduction with at 5% significance level and with 15% type II error rate (85% power)?

Solution

- $z_\alpha = z_{0.05} = 1.64$



- $z_\beta = z_{0.15} = 1.03$



- $n = \left(\frac{(z_\alpha + z_\beta)\sigma}{\Delta}\right)^2 = \left(\frac{(1.64+1.03)20}{2}\right)^2 = 718.93 \to 719$

Chi-square Test for Independence

The test is applied when you have two categorical variables from a single population. It is used to determine whether there is a significant association between the two variables.

χ^2 test is applied to a contingency table with two factors

- H_0 : factors are independent
- H_a : factors are dependent

Problem 4.1

A restaurant owner surveys a random sample of 385 customers to determine whether customer satisfaction is related to gender and age.

	Young Male	Young Female	Adult Male	Adult Female
Satisfied	25	30	135	112
Not satisfied	8	16	22	37

Problem 4.1

A restaurant owner surveys a random sample of 385 customers to determine whether customer satisfaction is related to gender and age.

	Young Male	Young Female	Adult Male	Adult Female
Satisfied	25	30	135	112
Not satisfied	8	16	22	37

Solution

	Young M	Young F	Adult M	Adult F	Total
Satisfied	25	30	135	112	302
Not satisfied	8	16	22	37	83
Total	33	46	157	149	385

If gender/age and satisfaction were independent then $\mathrm{P}(\textit{satisfied} \cap \textit{young male}) = \mathrm{P}(\textit{satisfied})\,\mathrm{P}(\textit{young male})$

Observed and Expected

- $P(\text{satisfied}) = 302/385$
- $P(\text{young male}) = 33/385$
- $\text{P(satisfied} \cap \text{young male)} = 302 * 33/385^2$
- Expected number of satisfied young males $= 302 * 33/385$

Observed:

	Young M	Young F	Adult M	Adult F	Total
Satisfied	25	30	135	112	302
Not satisfied	8	16	22	37	83
Total	33	46	157	149	385

Expected:

	Young M	Young F	Adult M	Adult F	Total
Satisfied	25.9	36.1	123.1	116.9	302
Not satisfied	7.1	9.9	33.9	32.1	83
Total	33	46	157	149	385

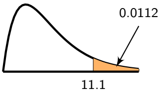
Chi-square Test for Independence

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

$$\chi^2 = \frac{(25-25.9)^2}{25.9} + \frac{(30-36.1)^2}{36.1} + \ldots = 11.1$$

$$df = (n-1)(m-1) = (2-1)(4-1) = 3$$

$$\mathrm{P}(\chi^2(3) \geq 11.1) = 0.112$$



At 5% significance level H_0 is rejected, i.e., there is evidence in this data that gender/age and satisfaction are not independent.

Chi-square Goodness of Fit

Problem 4.2

A grocery store manager wishes to determine whether a certain product will sell equally well in any of the five locations in the store. Five displays are set up, one for each location, and the resulting numbers of the product sold are noted

Location	*1*	*2*	*3*	*4*	*5*
Items sold	*43*	*29*	*52*	*34*	*48*

Is there enough evidence to claim a difference?

Chi-square Goodness of Fit

Problem 4.2

A grocery store manager wishes to determine whether a certain product will sell equally well in any of the five locations in the store. Five displays are set up, one for each location, and the resulting numbers of the product sold are noted

Location	*1*	*2*	*3*	*4*	*5*
Items sold	*43*	*29*	*52*	*34*	*48*

Is there enough evidence to claim a difference?

Solution

- H_0 : *The distribution is uniform*

Chi-square Goodness of Fit

Problem 4.2

A grocery store manager wishes to determine whether a certain product will sell equally well in any of the five locations in the store. Five displays are set up, one for each location, and the resulting numbers of the product sold are noted

Location	1	2	3	4	5
Items sold	43	29	52	34	48

Is there enough evidence to claim a difference?

Solution

- H_0 : *The distribution is uniform*
- H_a : *The distribution is not uniform*

Chi-square Goodness of Fit

Problem 4.2

A grocery store manager wishes to determine whether a certain product will sell equally well in any of the five locations in the store. Five displays are set up, one for each location, and the resulting numbers of the product sold are noted

Location	1	2	3	4	5
Items sold	43	29	52	34	48

Is there enough evidence to claim a difference?

Solution

- H_0 : *The distribution is uniform*
- H_a : *The distribution is not uniform*
- $Total = 43+29+52+34+48=206$

Chi-square Goodness of Fit

Problem 4.2

A grocery store manager wishes to determine whether a certain product will sell equally well in any of the five locations in the store. Five displays are set up, one for each location, and the resulting numbers of the product sold are noted

Location	*1*	*2*	*3*	*4*	*5*
Items sold	*43*	*29*	*52*	*34*	*48*

Is there enough evidence to claim a difference?

Solution

- H_0 : *The distribution is uniform*
- H_a : *The distribution is not uniform*
- *Total* $= 43+29+52+34+48=206$
- *We expect* $206/5=41.2$ *units sold in each location*

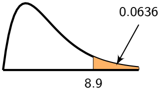
Chi-square Goodness of Fit

Location	1	2	3	4	5
Items sold	43	29	52	34	48
Expected	41.2	41.2	41.2	41.2	41.2

$$\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(43-41.2)^2}{41.2} + \ldots = 8.9$$

$$df = n - 1$$

$$\mathrm{P}(\chi^2(4) \geq 8.9) = 0.0636$$



At 5% significance level H_0 is not rejected, i.e., there is not enough evidence to claim that the five locations in the store are different.

Problem 4.3

A geneticist claims that four species of fruit flies should appear in the ratio of 1:3:3:9. Suppose that a sample of 4000 fruit flies contained 226, 764, 733, and 2277 flies of each species, respectively. At the 10% significance level, is there sufficient evidence to reject the geneticist's hypothesis?

Problem 4.3

A geneticist claims that four species of fruit flies should appear in the ratio of 1:3:3:9. Suppose that a sample of 4000 fruit flies contained 226, 764, 733, and 2277 flies of each species, respectively. At the 10% significance level, is there sufficient evidence to reject the geneticist's hypothesis?

Solution

- $\frac{1}{16} + \frac{3}{16} + \frac{3}{16} + \frac{9}{16} = 1,$ *that is* $4000 = 250 + 750 + 750 + 2250$

Observed	*226*	*764*	*733*	*2277*
Expected	*250*	*750*	*750*	*2250*

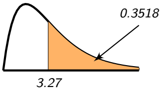
Problem 4.3

A geneticist claims that four species of fruit flies should appear in the ratio of 1:3:3:9. Suppose that a sample of 4000 fruit flies contained 226, 764, 733, and 2277 flies of each species, respectively. At the 10% significance level, is there sufficient evidence to reject the geneticist's hypothesis?

Solution

- $\frac{1}{16} + \frac{3}{16} + \frac{3}{16} + \frac{9}{16} = 1,$ *that is* $4000 = 250 + 750 + 750 + 2250$

Observed	*226*	*764*	*733*	*2277*
Expected	*250*	*750*	*750*	*2250*



- $\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(226-250)^2}{250} + \frac{(764-750)^2}{750} + \ldots = 3.27$

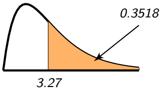
Problem 4.3

A geneticist claims that four species of fruit flies should appear in the ratio of 1:3:3:9. Suppose that a sample of 4000 fruit flies contained 226, 764, 733, and 2277 flies of each species, respectively. At the 10% significance level, is there sufficient evidence to reject the geneticist's hypothesis?

Solution

- $\frac{1}{16} + \frac{3}{16} + \frac{3}{16} + \frac{9}{16} = 1,$ *that is* $4000 = 250 + 750 + 750 + 2250$

Observed	*226*	*764*	*733*	*2277*
Expected	*250*	*750*	*750*	*2250*



- $\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(226-250)^2}{250} + \frac{(764-750)^2}{750} + \ldots = 3.27$
- *The geneticist's hypothesis about 1:3:3:9 ratio is not rejected at any reasonable significance level, there is no reason to believe it is not true.*

Problem 4.4

Weights of rice bags are supposed to have normal distribution. A random sample of 40 such bags was taken and the following frequencies were obtained.

weight	*below 480*	*480-490*	*490-500*	*500-510*	*510-520*	*above 520*
number of bags	*6*	*9*	*10*	*8*	*4*	*3*

Test the hypothesis that rice bags were chosen from a normal distribution with the mean weight of 500 *grams and standard deviation of* 18 *grams.*

Problem 4.4

Weights of rice bags are supposed to have normal distribution. A random sample of 40 such bags was taken and the following frequencies were obtained.

weight	below 480	480-490	490-500	500-510	510-520	above 520
number of bags	6	9	10	8	4	3

Test the hypothesis that rice bags were chosen from a normal distribution with the mean weight of 500 *grams and standard deviation of* 18 *grams.*

Solution

weight	below 480	480-490	490-500	500-510	510-520	above 520
	$z < -1.11$	$z \in (-1.11, -0.55]$	$z \in (-0.55, 0]$	$z \in (0, 0.55]$	$z \in (0.55, 1.11]$	$z > 1.11$
exp. prob	0.1333	0.156	0.2107	0.2107	0.156	0.1333
exp. count	5.3	6.2	8.4	8.4	6.2	5.3
observed	6	9	10	8	4	3

$$\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(6-5.3)^2}{5.3} + \cdots = 3.44$$

$P(\chi^2(5) > 3.44) = 0.63$, *i.e., there is no evidence against the claim that rice bags were chosen from a normal distribution with the mean weight of* 500 *grams and standard deviation of* 18 *grams.*

Chi-square test: Warning

- Chi-square test is applicable only if the expected value in each cell is greater than 5 (Compare to Binomial Distribution)

- Small expected values lead to higher uncertainty in $\chi^2 = \sum \frac{(O-E)^2}{E}$

- You might find Fisher exact test (Hypergeometric test) also useful

Hypergeometric Test

Problem 4.5

A sample of teenagers might be divided into male and female on the one hand, and those that are and are not currently dieting on the other. We hypothesize, perhaps, that the proportion of dieting individuals is higher among the women than among the men, and we want to test whether any difference of proportions that we observe is significant.

	Men	Women	Total
Dieting	1	9	10
Not dieting	11	3	14
Total	12	12	24

Hypergeometric Test

Problem 4.5

A sample of teenagers might be divided into male and female on the one hand, and those that are and are not currently dieting on the other. We hypothesize, perhaps, that the proportion of dieting individuals is higher among the women than among the men, and we want to test whether any difference of proportions that we observe is significant.

	Men	*Women*	*Total*
Dieting	*1*	*9*	*10*
Not dieting	*11*	*3*	*14*
Total	*12*	*12*	*24*

Solution

	Men	*Women*	*Total*
Dieting	*5*	*5*	*10*
Not dieting	*7*	*7*	*14*
Total	*12*	*12*	*24*

Expected < 5

Hypergeometric Test

	Men	Women	Total
Dieting	1	9	10
Not dieting	11	3	14
Total	12	12	24

	Men	Women	Total
Dieting	a	b	$a+b$
Not dieting	c	d	$c+d$
Total	$a+c$	$b+d$	n

$$P = \frac{\binom{a+b}{a}\binom{c+d}{c}}{\binom{n}{a+c}} = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{n!a!b!c!d!}$$

$$P = \frac{10!14!12!12!}{24!1!9!11!3!} = 0.0013$$

Note that

- Exact computation with factorials of large numbers is troublesome
- Hypergeometric test is a **point** test, i.e., it estimates the probability of **exactly** the table that was observed. If you are interested in deviations in certain direction, you have to repeat hypergeometric test to compute hypergeometric CDF

Sign test

The sign test is a method to find consistent *ordinal* differences between pairs of observations. It determines if one member in the pair of observations tends to be greater than the other member. Unlike t-test, there is no assumption of normality for small samples, neither any other assumption about the nature of the random variable.

- H_0 : $\text{median}_1 = \text{median}_2$
- H_a : $\text{median}_1 > \text{median}_2$

Sample (X_i, Y_i), $i = 1 \ldots n$
$\hat{p} =$ sample proportion of $X_i > Y_i$
Ties are split randomly between $X_i > Y_i$ and $X_i < Y_i$

Sign test

Problem 4.6

The following data was collected about the weights of ten patients in the treatment group taking certain weight-control medication. Do these data suggest that the weight-control medication works?

Patient	Before	After
1	200	197
2	202	204
3	194	167
4	188	192
5	166	166

Patient	Before	After
6	196	190
7	180	176
8	188	182
9	180	180
10	210	202

Sign test

Problem 4.6

The following data was collected about the weights of ten patients in the treatment group taking certain weight-control medication. Do these data suggest that the weight-control medication works?

Patient	Before	After
1	200	197
2	202	204
3	194	167
4	188	192
5	166	166

Patient	Before	After
6	196	190
7	180	176
8	188	182
9	180	180
10	210	202

Solution

- *Out of 10 patients, 5 reduced weight, 3 gained weight, and 2 stayed unchanged.*

Sign test

Problem 4.6

The following data was collected about the weights of ten patients in the treatment group taking certain weight-control medication. Do these data suggest that the weight-control medication works?

Patient	Before	After
1	200	197
2	202	204
3	194	167
4	188	192
5	166	166

Patient	Before	After
6	196	190
7	180	176
8	188	182
9	180	180
10	210	202

Solution

- *Out of 10 patients, 5 reduced weight, 3 gained weight, and 2 stayed unchanged.*
- $X \sim Bi(n = 10, p = 0.5)$

Sign test

Problem 4.6

The following data was collected about the weights of ten patients in the treatment group taking certain weight-control medication. Do these data suggest that the weight-control medication works?

Patient	Before	After	Patient	Before	After
1	200	197	6	196	190
2	202	204	7	180	176
3	194	167	8	188	182
4	188	192	9	180	180
5	166	166	10	210	202

Solution

- *Out of 10 patients, 5 reduced weight, 3 gained weight, and 2 stayed unchanged.*
- $X \sim Bi(n = 10, p = 0.5)$
- $\mathrm{P}(X \geq 6) = \mathrm{P}(X = 6) + \mathrm{P}(X = 7) + \cdots + \mathrm{P}(X = 10) = 0.3770$, *there is not enough evidence to claim that the medication works.*

Mann-Whitney U-test

Wilcoxon-Mann-Whitney test

- X and Y are two populations

Mann-Whitney U-test

Wilcoxon-Mann-Whitney test

- X and Y are two populations
- $H_0 : \mathrm{P}(X > Y) = \mathrm{P}(Y > X)$

Mann-Whitney U-test

Wilcoxon-Mann-Whitney test

- X and Y are two populations
- $H_0 : \mathrm{P}(X > Y) = \mathrm{P}(Y > X)$
- $H_a : \mathrm{P}(X > Y) \neq \mathrm{P}(Y > X)$

Mann-Whitney U-test

Wilcoxon-Mann-Whitney test

- X and Y are two populations
- $H_0 : \mathrm{P}(X > Y) = \mathrm{P}(Y > X)$
- $H_a : \mathrm{P}(X > Y) \neq \mathrm{P}(Y > X)$
- U-statistic

Mann-Whitney U-test

Wilcoxon-Mann-Whitney test

- X and Y are two populations
- $H_0 : \mathrm{P}(X > Y) = \mathrm{P}(Y > X)$
- $H_a : \mathrm{P}(X > Y) \neq \mathrm{P}(Y > X)$
- U-statistic
  - $\{X_1, \ldots, X_n\}$ and $\{Y_1, \ldots, Y_m\}$ are two samples

Mann-Whitney U-test

Wilcoxon-Mann-Whitney test

- X and Y are two populations
- $H_0 : \mathrm{P}(X > Y) = \mathrm{P}(Y > X)$
- $H_a : \mathrm{P}(X > Y) \neq \mathrm{P}(Y > X)$
- U-statistic
  - $\{X_1, \ldots, X_n\}$ and $\{Y_1, \ldots, Y_m\}$ are two samples
  - Assign ranks to all the observations $\{X_1, \ldots, X_n, Y_1, \ldots, Y_m\}$

Mann-Whitney U-test

Wilcoxon-Mann-Whitney test

- X and Y are two populations
- $H_0 : \mathrm{P}(X > Y) = \mathrm{P}(Y > X)$
- $H_a : \mathrm{P}(X > Y) \neq \mathrm{P}(Y > X)$
- U-statistic
  - $\{X_1, \ldots, X_n\}$ and $\{Y_1, \ldots, Y_m\}$ are two samples
  - Assign ranks to all the observations $\{X_1, \ldots, X_n, Y_1, \ldots, Y_m\}$
  - R_1 = the sum of ranks for the observations which came from sample 1

Mann-Whitney U-test

Wilcoxon-Mann-Whitney test

- X and Y are two populations
- $H_0 : \mathrm{P}(X > Y) = \mathrm{P}(Y > X)$
- $H_a : \mathrm{P}(X > Y) \neq \mathrm{P}(Y > X)$
- U-statistic
    - $\{X_1, \dots, X_n\}$ and $\{Y_1, \dots, Y_m\}$ are two samples
    - Assign ranks to all the observations $\{X_1, \dots, X_n, Y_1, \dots, Y_m\}$
    - $R_1 =$ the sum of ranks for the observations which came from sample 1
    - $R_2 =$ the sum of ranks for the observations which came from sample 2

Mann-Whitney U-test

Wilcoxon-Mann-Whitney test

- X and Y are two populations
- $H_0 : \mathrm{P}(X > Y) = \mathrm{P}(Y > X)$
- $H_a : \mathrm{P}(X > Y) \neq \mathrm{P}(Y > X)$
- U-statistic
  - $\{X_1, \ldots, X_n\}$ and $\{Y_1, \ldots, Y_m\}$ are two samples
  - Assign ranks to all the observations $\{X_1, \ldots, X_n, Y_1, \ldots, Y_m\}$
  - $R_1 =$ the sum of ranks for the observations which came from sample 1
  - $R_2 =$ the sum of ranks for the observations which came from sample 2
  - $U_1 = R_1 - \frac{n(n+1)}{2} \quad U_2 = R_2 - \frac{m(m+1)}{2}$

Mann-Whitney U-test

Wilcoxon-Mann-Whitney test

- X and Y are two populations
- $H_0 : \mathrm{P}(X > Y) = \mathrm{P}(Y > X)$
- $H_a : \mathrm{P}(X > Y) \neq \mathrm{P}(Y > X)$
- U-statistic
  - $\{X_1, \ldots, X_n\}$ and $\{Y_1, \ldots, Y_m\}$ are two samples
  - Assign ranks to all the observations $\{X_1, \ldots, X_n, Y_1, \ldots, Y_m\}$
  - $R_1 =$ the sum of ranks for the observations which came from sample 1
  - $R_2 =$ the sum of ranks for the observations which came from sample 2
  - $U_1 = R_1 - \frac{n(n+1)}{2} \quad U_2 = R_2 - \frac{m(m+1)}{2}$
  - $U = \max\{U_1, U_2\}$

Mann-Whitney U-test

Wilcoxon-Mann-Whitney test

- X and Y are two populations
- $H_0 : \mathrm{P}(X > Y) = \mathrm{P}(Y > X)$
- $H_a : \mathrm{P}(X > Y) \neq \mathrm{P}(Y > X)$
- U-statistic
  - $\{X_1, \ldots, X_n\}$ and $\{Y_1, \ldots, Y_m\}$ are two samples
  - Assign ranks to all the observations $\{X_1, \ldots, X_n, Y_1, \ldots, Y_m\}$
  - $R_1 =$ the sum of ranks for the observations which came from sample 1
  - $R_2 =$ the sum of ranks for the observations which came from sample 2
  - $U_1 = R_1 - \frac{n(n+1)}{2} \quad U_2 = R_2 - \frac{m(m+1)}{2}$
  - $U = \max\{U_1, U_2\}$
  - In case of ties there is a small correction to this procedure

Mann-Whitney critical values and probabilities

Critical values $p = 0.05$

$n_1 \backslash n_2$	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	1	2	2	3	3	4	4	5	5	6	6	7	7	8
4	3	4	4	5	6	7	8	9	10	11	11	12	13	13
5	5	6	7	8	9	11	12	13	14	15	17	18	19	20
6	6	8	10	11	13	14	16	17	19	21	22	24	25	27
7	8	10	12	14	16	18	20	22	24	26	28	30	32	34
8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
9	12	15	17	20	23	26	28	31	34	37	39	42	45	48
10	14	17	20	23	26	29	33	36	39	42	45	48	52	55
11	16	19	23	26	30	33	37	40	44	47	51	55	58	62
12	18	22	26	29	33	37	41	45	49	53	57	61	65	69

$$U \sim \mathcal{N}(\mu, \sigma)$$

$$\mu = \frac{n_1 n_2}{2}$$

$$\sigma = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

Problem 4.7

A hospital exercise laboratory technician notes the resting pulse rates of five joggers to be 60, 58, 59, 61, and 67, respectively, while the resting pulse rates of seven non-exercisers are 83, 60, 75, 71, 91, 82, and 84, respectively. Use Mann-Whitney criterion to test whether resting pulse rates of joggers tend to be different from the resting pulse rates of non-exercisers.

Problem 4.7

A hospital exercise laboratory technician notes the resting pulse rates of five joggers to be 60, 58, 59, 61, and 67, respectively, while the resting pulse rates of seven non-exercisers are 83, 60, 75, 71, 91, 82, and 84, respectively. Use Mann-Whitney criterion to test whether resting pulse rates of joggers tend to be different from the resting pulse rates of non-exercisers.

Solution

- $60, 58, 59, 61, 67, 83, 60, 75, 71, 91, 82, 84$

Problem 4.7

A hospital exercise laboratory technician notes the resting pulse rates of five joggers to be 60, 58, 59, 61, and 67, respectively, while the resting pulse rates of seven non-exercisers are 83, 60, 75, 71, 91, 82, and 84, respectively. Use Mann-Whitney criterion to test whether resting pulse rates of joggers tend to be different from the resting pulse rates of non-exercisers.

Solution

- $60, 58, 59, 61, 67, 83, 60, 75, 71, 91, 82, 84$
- $58, 59, 60, 60, 61, 67, 71, 75, 82, 83, 84, 91$

Problem 4.7

A hospital exercise laboratory technician notes the resting pulse rates of five joggers to be 60, 58, 59, 61, and 67, respectively, while the resting pulse rates of seven non-exercisers are 83, 60, 75, 71, 91, 82, and 84, respectively. Use Mann-Whitney criterion to test whether resting pulse rates of joggers tend to be different from the resting pulse rates of non-exercisers.

Solution

- $60, 58, 59, 61, 67, 83, 60, 75, 71, 91, 82, 84$
- $58, 59, 60, 60, 61, 67, 71, 75, 82, 83, 84, 91$
- $1, 2, 3.5, 3.5, 5, 6, 7, 8, 9, 10, 11, 12$

Problem 4.7

A hospital exercise laboratory technician notes the resting pulse rates of five joggers to be 60, 58, 59, 61, and 67, respectively, while the resting pulse rates of seven non-exercisers are 83, 60, 75, 71, 91, 82, and 84, respectively. Use Mann-Whitney criterion to test whether resting pulse rates of joggers tend to be different from the resting pulse rates of non-exercisers.

Solution

- $60, 58, 59, 61, 67, 83, 60, 75, 71, 91, 82, 84$
- $58, 59, 60, 60, 61, 67, 71, 75, 82, 83, 84, 91$
- $1, 2, 3.5, 3.5, 5, 6, 7, 8, 9, 10, 11, 12$
- $R_1 = 1 + 2 + 3.5 + 5 + 6, \ R_2 = 3.5 + 7 + 8 + 9 + 10 + 11 + 12$

Problem 4.7

A hospital exercise laboratory technician notes the resting pulse rates of five joggers to be 60, 58, 59, 61, and 67, respectively, while the resting pulse rates of seven non-exercisers are 83, 60, 75, 71, 91, 82, and 84, respectively. Use Mann-Whitney criterion to test whether resting pulse rates of joggers tend to be different from the resting pulse rates of non-exercisers.

Solution

- $60, 58, 59, 61, 67, 83, 60, 75, 71, 91, 82, 84$
- $58, 59, 60, 60, 61, 67, 71, 75, 82, 83, 84, 91$
- $1, 2, 3.5, 3.5, 5, 6, 7, 8, 9, 10, 11, 12$
- $R_1 = 1 + 2 + 3.5 + 5 + 6,\ R_2 = 3.5 + 7 + 8 + 9 + 10 + 11 + 12$
- $U_1 = 17.5 - 5 * 4/2 = 7.5,\ U_2 = 60.5 - 7 * 6/2 = 39.5$

Problem 4.7

A hospital exercise laboratory technician notes the resting pulse rates of five joggers to be 60, 58, 59, 61, and 67, respectively, while the resting pulse rates of seven non-exercisers are 83, 60, 75, 71, 91, 82, and 84, respectively. Use Mann-Whitney criterion to test whether resting pulse rates of joggers tend to be different from the resting pulse rates of non-exercisers.

Solution

- $60, 58, 59, 61, 67, 83, 60, 75, 71, 91, 82, 84$
- $58, 59, 60, 60, 61, 67, 71, 75, 82, 83, 84, 91$
- $1, 2, 3.5, 3.5, 5, 6, 7, 8, 9, 10, 11, 12$
- $R_1 = 1 + 2 + 3.5 + 5 + 6,\ R_2 = 3.5 + 7 + 8 + 9 + 10 + 11 + 12$
- $U_1 = 17.5 - 5 * 4/2 = 7.5,\ U_2 = 60.5 - 7 * 6/2 = 39.5$
- $U = 39.5 > U_{0.05}(5, 7) = 5$, *therefore* H_0 *is rejected, i.e. there is enough evidence at the* 5% *significance level that the resting pulse rates of joggers are different from the resting pulse rates of non-exercisers.*

Wilcoxon signed-rank test

The Wilcoxon signed-rank test is used to assess whether the differences are symmetric and centered around zero

- H_0 : differences follow a symmetric distribution around zero

Wilcoxon signed-rank test

The Wilcoxon signed-rank test is used to assess whether the differences are symmetric and centered around zero

- H_0 : differences follow a symmetric distribution around zero
- H_1 : differences don't follow a symmetric distribution around zero

Wilcoxon signed-rank test

The Wilcoxon signed-rank test is used to assess whether the differences are symmetric and centered around zero

- H_0 : differences follow a symmetric distribution around zero
- H_1 : differences don't follow a symmetric distribution around zero
- W-statistic

Wilcoxon signed-rank test

The Wilcoxon signed-rank test is used to assess whether the differences are symmetric and centered around zero

- H_0 : differences follow a symmetric distribution around zero
- H_1 : differences don't follow a symmetric distribution around zero
- W-statistic
    - $\{X_1, \ldots, X_n\}$ and $\{Y_1, \ldots, Y_n\}$ are *paired* samples

Wilcoxon signed-rank test

The Wilcoxon signed-rank test is used to assess whether the differences are symmetric and centered around zero

- H_0 : differences follow a symmetric distribution around zero
- H_1 : differences don't follow a symmetric distribution around zero
- W-statistic
    - $\{X_1, \ldots, X_n\}$ and $\{Y_1, \ldots, Y_n\}$ are *paired* samples
    - Compute $d_i = |X_i - Y_i| = 0$ and exclude elements with $d_i = 0$

Wilcoxon signed-rank test

The Wilcoxon signed-rank test is used to assess whether the differences are symmetric and centered around zero

- H_0 : differences follow a symmetric distribution around zero
- H_1 : differences don't follow a symmetric distribution around zero
- W-statistic
  - $\{X_1, \ldots, X_n\}$ and $\{Y_1, \ldots, Y_n\}$ are *paired* samples
  - Compute $d_i = |X_i - Y_i| = 0$ and exclude elements with $d_i = 0$
  - Sort d_i ascending

Wilcoxon signed-rank test

The Wilcoxon signed-rank test is used to assess whether the differences are symmetric and centered around zero

- H_0 : differences follow a symmetric distribution around zero
- H_1 : differences don't follow a symmetric distribution around zero
- W-statistic
  - $\{X_1, \ldots, X_n\}$ and $\{Y_1, \ldots, Y_n\}$ are *paired* samples
  - Compute $d_i = |X_i - Y_i| = 0$ and exclude elements with $d_i = 0$
  - Sort d_i ascending
  - $W = \sum sgn(X_i - Y_i) * R_i$, where R_i is the rank of d_i

Wilcoxon signed-rank test

The Wilcoxon signed-rank test is used to assess whether the differences are symmetric and centered around zero

- H_0 : differences follow a symmetric distribution around zero
- H_1 : differences don't follow a symmetric distribution around zero
- W-statistic
  - $\{X_1, \ldots, X_n\}$ and $\{Y_1, \ldots, Y_n\}$ are *paired* samples
  - Compute $d_i = |X_i - Y_i| = 0$ and exclude elements with $d_i = 0$
  - Sort d_i ascending
  - $W = \sum sgn(X_i - Y_i) * R_i$, where R_i is the rank of d_i
  - $W \sim N\left(\mu = 0, \sigma = \sqrt{\frac{n(n+1)(2n+1)}{6}}\right)$ for $n \geq 10$

Problem 4.8

Twelve volunteers tested the efficacy of a new fuel additive in their cars. They first ride a full tank without additive and record the number of miles to reach the fuel indicator threshold, and then re-fuel with the additive and repeat the same measurement until the indicator light shows on. The following data were obtained without: 125.3, 101.0, 117.2, 133.7, 96.4, 124.5, 118.7, 106.2, 116.3, 120.2, 125.0, 128.8, and with additive 127.3, 120.2, 126.2, 125.4, 115.1, 118.5, 135.5, 118.2, 122.9, 120.1, 120.8, 130.7

Problem 4.8

Twelve volunteers tested the efficacy of a new fuel additive in their cars. They first ride a full tank without additive and record the number of miles to reach the fuel indicator threshold, and then re-fuel with the additive and repeat the same measurement until the indicator light shows on. The following data were obtained without: 125.3, 101.0, 117.2, 133.7, 96.4, 124.5, 118.7, 106.2, 116.3, 120.2, 125.0, 128.8, and with additive 127.3, 120.2, 126.2, 125.4, 115.1, 118.5, 135.5, 118.2, 122.9, 120.1, 120.8, 130.7

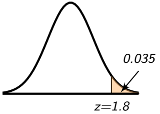
Solution

N	before	after	d	|d|	sign	rank	sign * rank
1	125.3	127.3	2	2	1	3	3
2	101	120.2	19.2	19.2	1	12	12
3	117.2	126.2	9	9	1	8	8
4	133.7	125.4	-8.3	8.3	-1	7	-7
5	96.4	115.1	18.7	18.7	1	11	11
6	124.5	118.5	-6	6	-1	5	-5
7	118.7	135.5	16.8	16.8	1	10	10
8	106.2	118.2	12	12	1	9	9
9	116.3	122.9	6.6	6.6	1	6	6
10	120.2	120.1	0.1	0.1	1	1	1
11	125	120.8	-4.2	4.2	-1	4	-4
12	128.8	130.7	1.9	1.9	1	2	2

$$W = 46, n = 11$$

$$\sigma = \sqrt{\frac{n(n+1)(2n+1)}{6}} = 25.5$$

$$z = \frac{46 - 0}{25.5} = 1.80$$

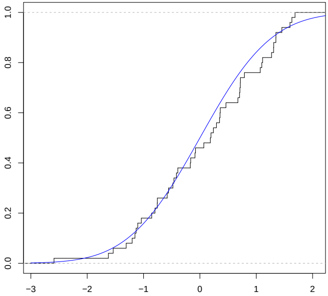


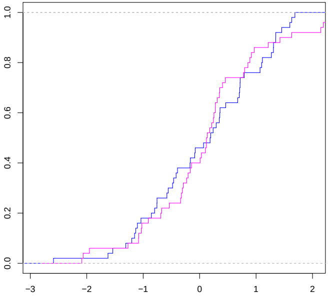
H_0 is rejected at the 5% sign. level

Kolmogorov-Smirnov test

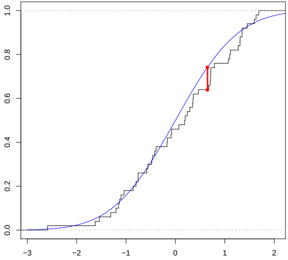
The Kolmogorov-Smirnov test (KS test) is a non-parametric test to check whether the empirical cumulative distribution function (eCDF) comes from a reference probability distribution, or whether two eCDFs come from the same reference distribution.

- eCDF comes from a reference probability distribution (one-sample KS test)
- two eCDFs come from the same reference distribution (two-sample KS test)



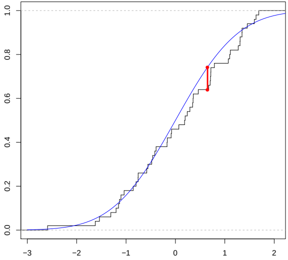


One-sample Kolmogorov-Smirnov test



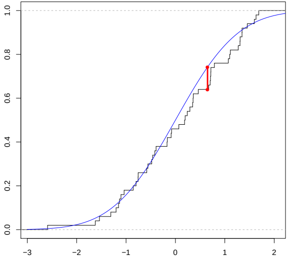
- $D_n = \sup_x |F_n(x) - F(x)|$

One-sample Kolmogorov-Smirnov test



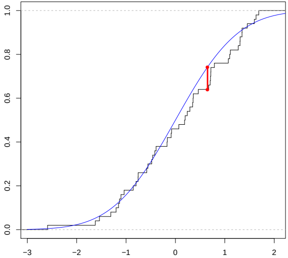
- $D_n = \sup_x |F_n(x) - F(x)|$
- If $F(x)$ is continuous then D_n doesn't depend on $F(x)$

One-sample Kolmogorov-Smirnov test



- $D_n = \sup_x |F_n(x) - F(x)|$
- If $F(x)$ is continuous then D_n doesn't depend on $F(x)$
- $\mathrm{P}(\sqrt{n}D_n \leq x) = H(x) = 1 - 2\sum_{k=1}^{\infty}(-1)^{k-1}e^{-2k^2x}$

One-sample Kolmogorov-Smirnov test



- $D_n = \sup_x |F_n(x) - F(x)|$
- If $F(x)$ is continuous then D_n doesn't depend on $F(x)$
- $P(\sqrt{n}D_n \leq x) = H(x) = 1 - 2\sum_{k=1}^{\infty}(-1)^{k-1}e^{-2k^2x}$
- $H(x)$ is called Kolmogorov-Smirnov distribution

Kolmogorov-Smirnov distribution

Critical values for $\sup_x |F_n(x) - F(x)|$

	Level of significance, α			
n	0.10	0.05	0.02	0.01
1	0.95000	0.97500	0.99000	0.99500
2	0.77639	0.84189	0.90000	0.92929
3	0.63604	0.70760	0.78456	0.82900
4	0.56522	0.62394	0.68887	0.73424
5	0.50945	0.56328	0.62718	0.66853
6	0.46799	0.51926	0.57741	0.61661
7	0.43607	0.48342	0.53844	0.57581
8	0.40962	0.45427	0.50654	0.54179
9	0.38746	0.43001	0.47960	0.51332
10	0.36866	0.40925	0.45662	0.48893
11	0.35242	0.39122	0.43670	0.46770
12	0.33815	0.37543	0.41918	0.44905
13	0.32549	0.36143	0.40362	0.43247
14	0.31417	0.34890	0.38970	0.41762
15	0.30397	0.33760	0.37713	0.40420
16	0.29472	0.32733	0.36571	0.39201
17	0.28627	0.31796	0.35528	0.38086
18	0.27851	0.30936	0.34569	0.37062
19	0.27136	0.30143	0.33685	0.36117
20	0.26473	0.29408	0.32866	0.35241

Problem 4.9

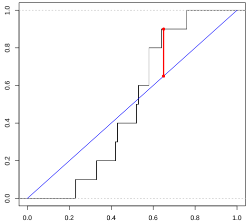
Test at the 5% significance level that the sample 0.58, 0.42, 0.52, 0.33, 0.43, 0.23, 0.58, 0.76, 0.53, 0.64 comes from a uniform distribution on the interval $[0, 1]$

Problem 4.9

Test at the 5% significance level that the sample 0.58, 0.42, 0.52, 0.33, 0.43, 0.23, 0.58, 0.76, 0.53, 0.64 comes from a uniform distribution on the interval $[0, 1]$

Solution

	x	*eCDF*
1	*0.23*	*0.10*
2	*0.33*	*0.20*
3	*0.42*	*0.30*
4	*0.43*	*0.40*
5	*0.52*	*0.50*
6	*0.53*	*0.60*
7	*0.58*	*0.70*
8	*0.58*	*0.80*
9	*0.64*	*0.90*
10	*0.76*	*1.00*

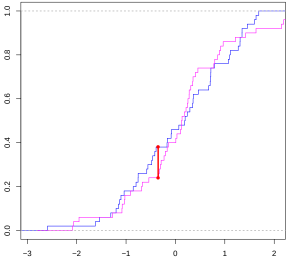


$$D = 0.90 - 0.64 = 0.26$$

$$\text{P}(D_n > 0.40925) = 0.05$$

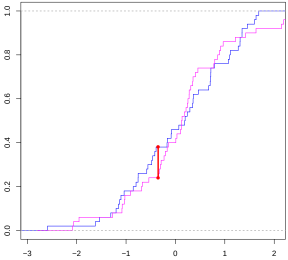
$0.26 < 0.40925$, *i.e., there is not enough evidence to reject* H_0, *i.e., it's not unlikely that the sample comes from a uniform distribution.*

Two-sample Kolmogorov-Smirnov test



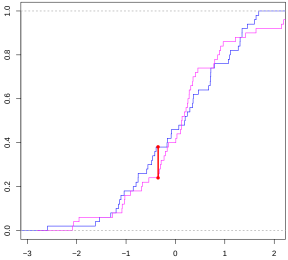
- H_0 : the two CDFs, $F_n(x)$ and $G_m(x)$, came from the same distribution

Two-sample Kolmogorov-Smirnov test



- H_0 : the two CDFs, $F_n(x)$ and $G_m(x)$, came from the same distribution
- $D_{n,m} = \sup_x |F_n(x) - G_m(x)|$

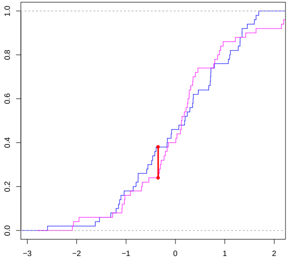
Two-sample Kolmogorov-Smirnov test



- H_0 : the two CDFs, $F_n(x)$ and $G_m(x)$, came from the same distribution
- $D_{n,m} = \sup_x |F_n(x) - G_m(x)|$
- $D_{n,m,\alpha} = c(\alpha)\sqrt{\frac{1}{n} + \frac{1}{m}}$

α	0.10	0.05	0.025	0.01	0.005	0.001
$c(\alpha)$	1.22	1.36	1.48	1.63	1.73	1.95

Two-sample Kolmogorov-Smirnov test



- H_0 : the two CDFs, $F_n(x)$ and $G_m(x)$, came from the same distribution
- $D_{n,m} = \sup_x |F_n(x) - G_m(x)|$
- $D_{n,m,\alpha} = c(\alpha)\sqrt{\frac{1}{n} + \frac{1}{m}}$

α	0.10	0.05	0.025	0.01	0.005	0.001
$c(\alpha)$	1.22	1.36	1.48	1.63	1.73	1.95

- Reject H_0 if $D_{n,m} > D_{n,m,\alpha}$

Problem 4.10

Test whether the following two samples come from the same distribution.

Sample 1: 0.05, 0.93, 0.62, 0.9, 0.84, 0.36, 0.26, 0.56, 0.02, 0.84

Sample 2: 0.39, 0.91, 0.86, 0.21, 0.39, 0.9, 0.1, 0.28, 0.02, 0.53, 0.08, 0.19

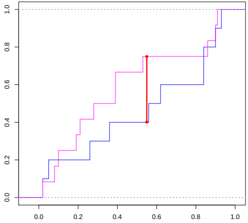
Problem 4.10

Test whether the following two samples come from the same distribution.

Sample 1: 0.05, 0.93, 0.62, 0.9, 0.84, 0.36, 0.26, 0.56, 0.02, 0.84

Sample 2: 0.39, 0.91, 0.86, 0.21, 0.39, 0.9, 0.1, 0.28, 0.02, 0.53, 0.08, 0.19

Solution



$$D = 0.35$$

$$n = 10, m = 12$$

$$D_{10,12,0.05} = 1.36 * \sqrt{\frac{1}{10} + \frac{1}{12}} = 0.58$$

$0.35 < 0.58$, *i.e., there is not enough evidence to reject* H_0, *i.e., it's not unlikely that the two samples come from different distributions.*

QQ-plot

A **QQ-plot** is a graphical method for comparing two probability distributions by plotting their quantiles against each other.

- $X = \{X_1, X_2, \ldots, X_n\} \to$ sorted: $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$

QQ-plot

A **QQ-plot** is a graphical method for comparing two probability distributions by plotting their quantiles against each other.

- $X = \{X_1, X_2, \ldots, X_n\} \rightarrow$ sorted: $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$
- $X_{(i)}$ – ith order statistic, i.e., the ith element in the ordered sample

QQ-plot

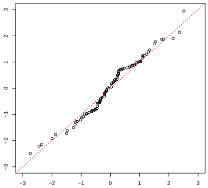
A **QQ-plot** is a graphical method for comparing two probability distributions by plotting their quantiles against each other.

- $X = \{X_1, X_2, \ldots, X_n\} \rightarrow$ sorted: $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$
- $X_{(i)}$ – ith order statistic, i.e., the ith element in the ordered sample
- $Y = \{Y_1, Y_2, \ldots, Y_n\} \rightarrow$ sorted: $Y_{(1)} \leq Y_{(2)} \leq \ldots \leq Y_{(n)}$

QQ-plot

A **QQ-plot** is a graphical method for comparing two probability distributions by plotting their quantiles against each other.

- $X = \{X_1, X_2, \ldots, X_n\} \rightarrow$ sorted: $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$
- $X_{(i)}$ – ith order statistic, i.e., the ith element in the ordered sample
- $Y = \{Y_1, Y_2, \ldots, Y_n\} \rightarrow$ sorted: $Y_{(1)} \leq Y_{(2)} \leq \ldots \leq Y_{(n)}$
- Plot $X_{(i)}$ vs $Y_{(i)}$



QQ-plot

- More generally plot sample **quantiles** against each other, or plot sample quantiles versus **theoretical quantiles**

QQ-plot

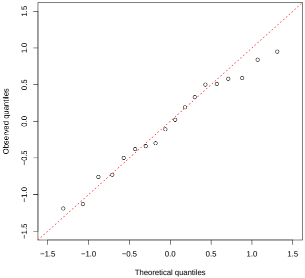
- More generally plot sample **quantiles** against each other, or plot sample quantiles versus **theoretical quantiles**
- Sorted sample: -1.26, -1.19, -1.13, -0.76, -0.73, -0.5, -0.38, -0.34, -0.3, -0.11, 0.02, 0.19, 0.33, 0.5, 0.51, 0.58, 0.59, 0.84, 0.95, 1

QQ-plot

- More generally plot sample **quantiles** against each other, or plot sample quantiles versus **theoretical quantiles**
- Sorted sample: -1.26, -1.19, -1.13, -0.76, -0.73, -0.5, -0.38, -0.34, -0.3, -0.11, 0.02, 0.19, 0.33, 0.5, 0.51, 0.58, 0.59, 0.84, 0.95, 1
- Probabilities equally spaced from 0 to 1: 0.05, 0.1, 0.14, 0.19, 0.24, 0.29, 0.33, 0.38, 0.43, 0.48, 0.52, 0.57, 0.62, 0.67, 0.71, 0.76, 0.81, 0.86, 0.9, 0.95

QQ-plot

- More generally plot sample **quantiles** against each other, or plot sample quantiles versus **theoretical quantiles**
- Sorted sample: -1.26, -1.19, -1.13, -0.76, -0.73, -0.5, -0.38, -0.34, -0.3, -0.11, 0.02, 0.19, 0.33, 0.5, 0.51, 0.58, 0.59, 0.84, 0.95, 1
- Probabilities equally spaced from 0 to 1: 0.05, 0.1, 0.14, 0.19, 0.24, 0.29, 0.33, 0.38, 0.43, 0.48, 0.52, 0.57, 0.62, 0.67, 0.71, 0.76, 0.81, 0.86, 0.9, 0.95
- Quantiles of the normal distribution: -1.67, -1.31, -1.07, -0.88, -0.71, -0.57, -0.43, -0.3, -0.18, -0.06, 0.06, 0.18, 0.3, 0.43, 0.57, 0.71, 0.88, 1.07, 1.31, 1.67



Shapiro normality test

The **Shapiro-Wilk** test checks whether a sample $X_1, ..., X_n$ came from a normally distributed population.

$$W = \frac{\sum_{i=1}^{n} a_i X_{(i)}}{\sum_{i=1}^{n} (X_{(i)} - \overline{X})^2}$$

- $X_{(i)}$ are ith order statistic, i.e., the ith element in the ordered sample

- Coefficients a_i are computed from the expected values of the order statistics of independent and identically distributed random variables sampled from the standard normal distribution, and from the covariance matrix of those order statistics

Correction for multiple testing

- As more symptoms are considered when testing the drug, it becomes more likely that it will do an improvement of at least one symptom

- As more types of side effects are considered when testing the drug, it becomes more likely that it will appear to be less safe in terms of at least one side effect

Familywise error rate

FWER is the probability of making **one or more** type I errors when performing multiple hypotheses tests

$$\tilde{\alpha} = 1 - (1 - \alpha)^k$$

- Bonferroni correction: use $\tilde{\alpha}/k$ per comparison
  - or multiply the P-value by k
- Šidák correction: $1 - (1 - \tilde{\alpha})^{\frac{1}{k}}$ per comparison
  - or transform P-value as $1 - (1 - p)^k$
- Holm-Bonferroni method: use different thresholds per comparison
  - order P-values from lowest to highest $p_1, \ldots, p_m$
  - reject H_i if $p_i < \frac{\tilde{\alpha}}{k-i+1}$

Summary

- Sample variance s^2 is proportional to χ^2 if the population is normal

Summary

- Sample variance s^2 is proportional to χ^2 if the population is normal
- The ratio of sample variances of independent samples has Fisher F-distribution

Summary

- Sample variance s^2 is proportional to χ^2 if the population is normal
- The ratio of sample variances of independent samples has Fisher F-distribution
- If population proportion is unknown, use $p = \frac{1}{2}$ to estimate sample size

Summary

- Sample variance s^2 is proportional to χ^2 if the population is normal
- The ratio of sample variances of independent samples has Fisher F-distribution
- If population proportion is unknown, use $p = \frac{1}{2}$ to estimate sample size
- χ^2 test for independence vs. χ^2 test for goodness of fit

Summary

- Sample variance s^2 is proportional to χ^2 if the population is normal
- The ratio of sample variances of independent samples has Fisher F-distribution
- If population proportion is unknown, use $p = \frac{1}{2}$ to estimate sample size
- χ^2 test for independence vs. χ^2 test for goodness of fit
- χ^2 test cannot be used when expected counts < 5

Summary

- Sample variance s^2 is proportional to χ^2 if the population is normal
- The ratio of sample variances of independent samples has Fisher F-distribution
- If population proportion is unknown, use $p = \frac{1}{2}$ to estimate sample size
- χ^2 test for independence vs. χ^2 test for goodness of fit
- χ^2 test cannot be used when expected counts < 5
- Sign, Mann-Whitney, and Wilcoxon test do not require normal population

Summary

- Sample variance s^2 is proportional to χ^2 if the population is normal
- The ratio of sample variances of independent samples has Fisher F-distribution
- If population proportion is unknown, use $p = \frac{1}{2}$ to estimate sample size
- χ^2 test for independence vs. χ^2 test for goodness of fit
- χ^2 test cannot be used when expected counts < 5
- Sign, Mann-Whitney, and Wilcoxon test do not require normal population
- One-sided KS test checks if a sample comes from the given population

Summary

- Sample variance s^2 is proportional to χ^2 if the population is normal
- The ratio of sample variances of independent samples has Fisher F-distribution
- If population proportion is unknown, use $p = \frac{1}{2}$ to estimate sample size
- χ^2 test for independence vs. χ^2 test for goodness of fit
- χ^2 test cannot be used when expected counts < 5
- Sign, Mann-Whitney, and Wilcoxon test do not require normal population
- One-sided KS test checks if a sample comes from the given population
- Two-sided KS test checks if two samples come from the same population

Summary

- Sample variance s^2 is proportional to χ^2 if the population is normal
- The ratio of sample variances of independent samples has Fisher F-distribution
- If population proportion is unknown, use $p = \frac{1}{2}$ to estimate sample size
- χ^2 test for independence vs. χ^2 test for goodness of fit
- χ^2 test cannot be used when expected counts < 5
- Sign, Mann-Whitney, and Wilcoxon test do not require normal population
- One-sided KS test checks if a sample comes from the given population
- Two-sided KS test checks if two samples come from the same population
- Familywise error rate is the probability of making one or more type I errors when performing multiple hypotheses tests