VE401 Probabilistic Methods in Eng.
Solution Manual for RC 5

Chen Xiwen

April 12, 2020

University of Michigan–Shanghai Jiao Tong University Joint Institute

Assignment 4.8

Let $X_1, \ldots, X_n$ be a random sample of size n from a continuous distribution with median M. Let

$$X_{\min} = \min_{1 \le i \le n} X_i, \qquad X_{\max} = \max_{1 \le i \le n} X_i.$$

1. Show that

$$P[X_{\min} \le M \le X_{\max}] = 1 - \left(\frac{1}{2}\right)^{n-1}.$$

2. Suppose the sample data are arranged from smallest to largest, so that $X_1 \le X_2 \le \cdots \le X_n$. Calculate

$$P[X_{k+1} \le M \le X_{n-k}]$$

for $k = 0, \ldots, \lfloor n/2 \rfloor$.

Solution. By definition of median,

$$P[X_i \le M] = \frac{1}{2},$$

for $i = 1, \ldots, n$.

1. We know that

$$\begin{aligned} P[X_{\min} \le M \le X_{\max}] &= 1 - P[X_{\min} > M] - P[X_{\max} < M] \\ &= 1 - 2 \cdot \left(\frac{1}{2}\right)^n \\ &= 1 - \left(\frac{1}{2}\right)^{n-1}. \end{aligned}$$

2. Suppose we have a set of random variables Y_i, where $i = 1, \ldots, n$, denoting the random variables before reordering. Then each Y_i is a Bernoulli random variable, and

$$P[Y_i \le M] = \frac{1}{2}.$$

Therefore, denoting

$$S = \{Y_1, \ldots, Y_n\},$$

we have

$$\begin{aligned}
P[X_{k+1} \le M \le X_{n-k}] &= 1 - P[X_{k+1} > M] - P[X_{n-k} < M] \\
P[X_{k+1} > M] &= \sum_{i=0}^{k} P\left[|\{Y_j \in S : Y_j \le M\}| = i\right] \\
&= \sum_{i=0}^{k} \binom{n}{i} \left(\frac{1}{2}\right)^n \\
P[X_{n-k} < M] &= \sum_{i=0}^{k} P\left[|\{Y_j \in S : Y_j \ge M\}| = i\right] \\
&= \sum_{i=0}^{k} \binom{n}{i} \left(\frac{1}{2}\right)^n .
\end{aligned}$$

Therefore,

$$P[X_{k+1} \le M \le X_{n-k}] = 1 - \sum_{i=0}^{k} \binom{n}{i} \left(\frac{1}{2}\right)^{n-1} .$$

Exercise 1.

The mean diameter of a metal rod produced by a machine is supposed to be 2.30mm. If the mean is different from this value by more than 0.2mm, it should be reported and concerned. Suppose the diameter of the rod X follows a normal distribution with mean μ and variance σ^2. A random sample of size $n = 16$ is tested with following data.

2.55	2.11	2.76	2.19
2.94	2.41	2.27	1.98
2.72	2.26	2.65	2.68
2.15	2.33	2.32	1.88

1. Set up and perform a suitable hypothesis test procedure for the mean such that the test conclusion will have at most 5% chance of being in error.
   Solution. We test the hypotheses

   $$H_0 : \mu = 2.30, \qquad H_1 : |\mu - 2.30| \ge 0.2.$$

   The test statistic is given by

   $$T_{n-1} = \frac{\overline{X} - 2.30}{S/\sqrt{n}} \quad \Rightarrow \quad t_{n-1} = 1.158,$$

   and critical value $t_{0.025,15} = 2.131$. Therefore, we fail to reject H_0 at the significance level of 5%.

2. Find a confidence interval for μ and compare with the results from (1).
Solution. A two-sided 95% confidence interval for μ is given by

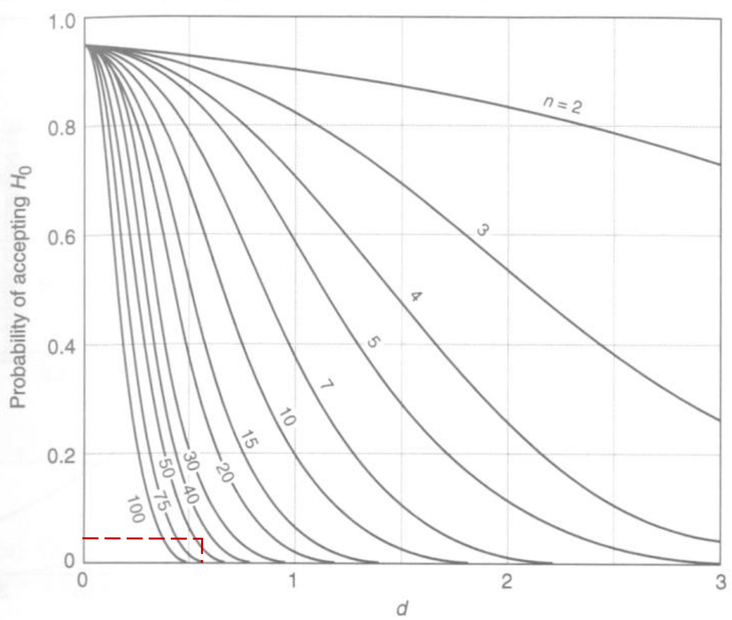
$$\overline{X} \pm \frac{t_{0.025,15}S}{\sqrt{n}} \quad \Rightarrow \quad \text{CI} = [2.227, 2.547].$$

The null value 2.30 falls inside this confidence interval. Therefore, H_0 is not rejected, which is in accordance with the conclusion in (1).

3. How large does the sample size should be to guarantee (1)?
Solution. We use the OC curve for two-sided T-test with $\alpha = 0.05$ and use the sample variance to estimate σ^2,

$$d = \frac{0.2}{S} \quad \Rightarrow \quad d \approx 0.66.$$

Chart VI Operating Characteristic Curves (*continued*)



(*e*) OC curves for different values of *n* for the two-sided *t* test for a level of significance $\alpha = 0.05$.

We read from OC curve that $n \approx 47$ would be a sufficient sample size.

4. If it is further required that the standard deviation of the normal distribution should be less than 0.5mm, set up and perform a suitable hypothesis test for the variance such that the test conclusion will have at most 5% chance of being in error.
Solution. We test the hypothesis

$$H_0 : \sigma \geq 0.5, \qquad H_1 : \sigma < 0.5$$

and obtain the test statistic

$$\chi^2_{n-1} = \frac{(n-1)S^2}{0.5^2} \quad \Rightarrow \quad \chi^2_{n-1} = 5.48,$$

and critical value $\chi^2_{0.95,15} = 7.261$. Therefore, we reject H_0 at significance level $\alpha = 5\%$ and thus accept H_1.

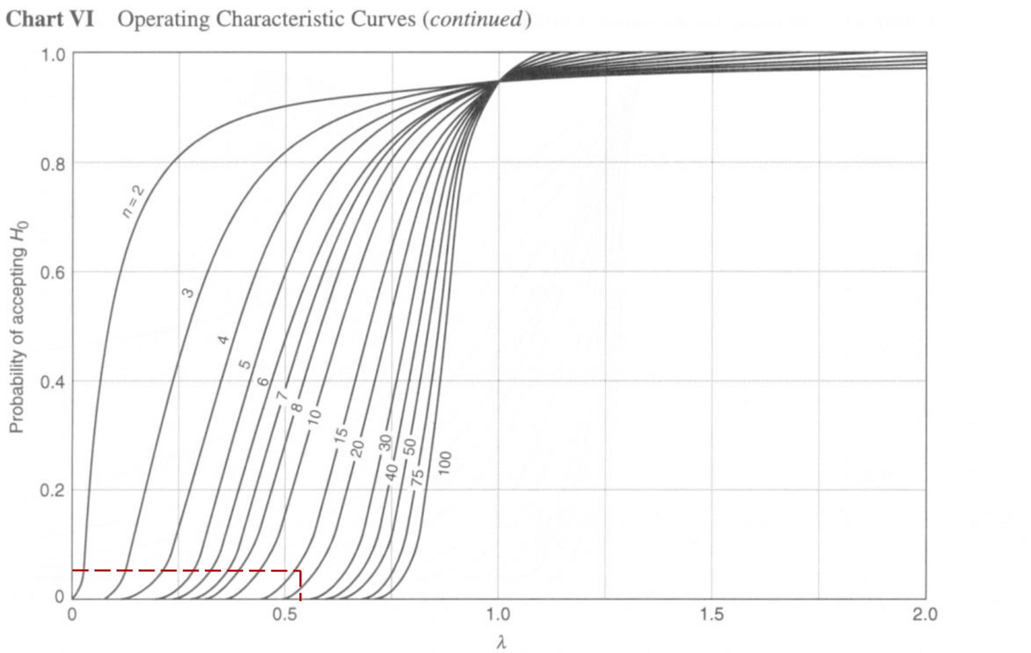
5. Find a confidence interval for σ^2 and compare with the results from (4).
Solution. A one-sided 95% confidence interval for σ^2 is given by

$$\sigma^2 < \frac{(n-1)S^2}{\chi_{0.95,n-1}} \quad \Rightarrow \quad \sigma^2 < 0.026,$$

and thus the null value $\sigma_0^2 = 0.25$ falls outside this interval. Therefore, H_0 is rejected, which is in accordance with the conclusion in (4).

6. What the true value of σ is necessary so that the test will have a power larger than 0.95?
Solution. We use the OC curve for one-sided Chi-squared test with $\alpha = 0.05$ and use the sample variance to estimate σ^2

$$\lambda = \frac{\sigma}{0.5} < 0.6 \quad \Rightarrow \quad \sigma < 0.3.$$



(m) OC curves for different values of n for the one-sided (lower tail) chi-square test for a level of significance $\alpha = 0.05$.

7. Suppose we have no knowledge about the distribution, perform a sign test for the following hypothesis for the median M

$$H_0 : M = 2.50$$

with $\alpha = 0.05$.
Solution. We get the following table of signs.

X_i	Sign	X_i	Sign	X_i	Sign	X_i	Sign
2.55	+	2.11	-	2.76	+	2.19	-
2.94	+	2.41	-	2.27	-	1.98	-
2.72	+	2.26	-	2.65	+	2.68	+
2.15	-	2 33	-	2.32	-	1.88	-

Therefore, $q_+ = 6, q_- = 10, n' = q_+ + q_- = 16$. Since it is a two-tailed test, we have the p-value

$$p\text{−value} = 2 \sum_{x=0}^{\min(q_+,q_-)} \binom{n'}{x} \frac{1}{2^n} = 0.454,$$

with which we fail to reject H_0.

8. Suppose we only know that the distribution of diameter is symmetric, perform a Wilcoxon signed rank test for the median M

$$H_0 : M = 2.50$$

with $\alpha = 0.05$.
Solution. We get the following table of signed ranks.

X_i	Rank	X_i	Rank	X_i	Rank	X_i	Rank
2.55	+1	2.68	+5	2.26	-9	2.11	-13
2.41	-2	2.32	-6	2.76	+10	2.94	+14
2.65	+3	2.72	+7	2.19	-11	1.98	-15
2.33	-4	2.27	-8	2.15	-12	1.88	-16

Then

$$W_+ = \sum_{R_i>0} R_i \quad \Rightarrow \quad w_+ = 40, \qquad |W_-| = \sum_{R_i<0} |R_i| \quad \Rightarrow \quad |w_-| = 96$$

with $n' = n = 16$. Looking up the table for <u>two-tailed</u> test with $\alpha = 0.05$, we get a critical value of 29, which is smaller than $\min(W_+, |W_-|) = 40$. Therefore, we fail to reject H_0. Alternatively, we approximate the distribution of $|W_-|$ by a normal distribution with

$$\mu = \frac{n(n+1)}{4} = 68, \qquad \sigma = \sqrt{\frac{n(n+1)(2n+1)}{24}} = 19.34$$

and normalize to obtain

$$z = \frac{|w_-| - \mu}{\sigma} = 1.45.$$

We have the critical value $z_{0.025} = 1.96 > z$. Therefore, we fail to reject H_0.

	alpha values						
n	0.001	0.005	0.01	0.025	0.05	0.10	0.20
5	--	--	--	--	--	0	2
6	--	--	--	--	0	2	3
7	--	--	--	0	2	3	5
8	--	--	0	2	3	5	8
9	--	0	1	3	5	8	10
10	--	1	3	5	8	10	14
11	0	3	5	8	10	13	17
12	1	5	7	10	13	17	21
13	2	7	9	13	17	21	26
14	4	9	12	17	21	25	31
15	6	12	15	20	25	30	36
16	8	15	19	25	29	35	42
17	11	19	23	29	34	41	48
18	14	23	27	34	40	47	55
19	18	27	32	39	46	53	62
20	21	32	37	45	52	60	69
21	25	37	42	51	58	67	77
22	30	42	48	57	65	75	86
23	35	48	54	64	73	83	94
24	40	54	61	72	81	91	104
25	45	60	68	79	89	100	113
26	51	67	75	87	98	110	124
27	57	74	83	96	107	119	134

	alpha values						
n	0.001	0.005	0.01	0.025	0.05	0.10	0.20
28	64	82	91	105	116	130	145
29	71	90	100	114	126	140	157
30	78	98	109	124	137	151	169
31	86	107	118	134	147	163	181
32	94	116	128	144	159	175	194
33	102	126	138	155	170	187	207
34	111	136	148	167	182	200	221
35	120	146	159	178	195	213	235
36	130	157	171	191	208	227	250
37	140	168	182	203	221	241	265
38	150	180	194	216	235	256	281
39	161	192	207	230	249	271	297
40	172	204	220	244	264	286	313
41	183	217	233	258	279	302	330
42	195	230	247	273	294	319	348
43	207	244	261	288	310	336	365
44	220	258	276	303	327	353	384
45	233	272	291	319	343	371	402
46	246	287	307	336	361	389	422
47	260	302	322	353	378	407	441
48	274	318	339	370	396	426	462
49	289	334	355	388	415	446	482
50	304	350	373	406	434	466	503

Note. The table above is for two-tailed tests. We need to double the α in one-tailed tests for table lookup, e.g., if $\alpha = 0.005$ is required, we use $\alpha = 0.01$ when looking up in the table.

9. It is said that if the diameter falls in a range of 2.30 ± 0.40mm, then the metal rod is of good quality. Perform the following test on proportion of good-quality rods that are produced by the machine.

$$H_0 : p \leq 0.50,$$

with $\alpha = 0.05$, where p is the proportion of good-quality rods.
Solution. We have the estimate for proportion

$$\widehat{p} = \frac{\#\{X_i : X_i \in (1.90, 2.70)\}}{n} = 0.75.$$

We have the statistic

$$Z = \frac{\widehat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} \quad \Rightarrow \quad z = 2.0,$$

and critical value $z_\alpha = 1.645 < z$. Therefore, we reject H_0 with at least 5% level of significance.

10. Another sample of 12 rods are tested. Perform a test comparing the proportions of good-quality rods of these two samples with $\alpha = 0.05$.

2.85	2.73	2.92	2.11
1.78	2.01	2.16	2.32
2.51	2.53	2.67	2.75

Solution. With the hypothesis

$$H_0 : p_1 = p_2,$$

we have the pooled proportion

$$\widehat{p} = \frac{n_1\widehat{p}_1 + n_2\widehat{p}_2}{n_1 + n_2} = 0.679,$$

where $\widehat{p}_2 = 0.583$. We have the test statistic

$$Z = \frac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\widehat{p}(1-\widehat{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} \quad \Rightarrow \quad z = 0.935$$

and critical value $z_{\alpha/2} = 1.96 > z$. Therefore, we fail to reject H_0.
Note. You might argue that the sample sizes of questions (9) and (10) are not sufficient for hypothesis tests for proportions. They are presented simply for illustrative purpose.

11. Perform a hypothesis test regarding the variances for the two samples with $\alpha = 0.05$.

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

Solution. We calculate

$$S_1^2 = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(X_i^{(1)} - \overline{X}_1)^2 \quad \Rightarrow \quad s_1^2 = 0.091$$

$$S_2^2 = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}(X_i^{(2)} - \overline{X}_2)^2 \quad \Rightarrow \quad s_2^2 = 0.133,$$

and thus the test statistics

$$F_{n_1-1,n_2-1} = \frac{S_1^2}{S_2^2} \Rightarrow f_{n_1-1,n_2-1} = 0.685, \qquad F_{n_2-1,n_1-1} = \frac{S_2^2}{S_1^2} \Rightarrow f_{n_2-1,n_1-1} = 1.460.$$

The critical values are given by

$$f_{0.025,15,11} = 3.33, \qquad f_{0.025,11,15} = 3.01.$$

Therefore, we fail to reject H_0.