Chapter 6   Statistics of the Kolmogorov-Smirnov Type

Section 6.1 The Kolmogorov goodness-of-fit test

The purpose of this section is to use the Kolmogorov goodness-of-fit test to check if the population distribution, from which a sample is drawn, is a specified distribution. The Kolmogorov goodness-of-fit test is more powerful than the χ^2 goodness-of-fit test.

- Kolmogorov Goodness-of-Fit Test
  (Kolmogorov-Smirnov One-Sample Test)

Let $x_1, x_2, \cdots, x_n$ be the observations of a random sample from a population with distribution function $F(.)$.

The empirical distribution function is defined as

$$S(x) = \begin{cases} 0 & x < x_{(1)} \\ 1/n & x_{(1)} \le x < x_{(2)} \\ 2/n & x_{(2)} \le x < x_{(3)} \\ & \vdots \\ (n-1)/n & x_{(n-1)} \le x < x_{(n)} \\ 1 & x \ge x_{(n)}. \end{cases}$$

$S(x) =$ fraction of x_i's that are less than or equal to x $(-\infty < x < \infty)$

Setting 1

$H_0 : F(x) \le F^*(x)$ for all $x \in (-\infty, \infty)$

$H_a : F(x) > F^*(x)$ for at least one value of x

Setting 2

$H_0 : F(x) \ge F^*(x)$ for all $x \in (-\infty, \infty)$

$H_a : F(x) < F^*(x)$ for at least one value of x

Setting 3

$H_0 : F(x) = F^*(x)$ for all $x \in (-\infty, \infty)$

$H_a : F(x) \ne F^*(x)$ for at least one value of x

Setting 1

$H_0 : F(x) \le F^*(x)$ for all $x \in (-\infty, \infty)$

$H_a : F(x) > F^*(x)$ for at least one value of x

Test statistic: $T^- = \sup_x \left(S(x) - F^*(x) \right)$

Reject H_0 if $T^-_{(\text{obs})} > w_{1-\alpha}$. (Table A13, one-sided test)

p-value $= P\left(T^- \ge T^-_{(\text{obs})}\right)$

Exact p-value $= t \sum_{j=0}^{[n(1-t)]} \binom{n}{j} \left(1 - t - \frac{j}{n}\right)^{n-j} \left(t + \frac{j}{n}\right)^{j-1} \quad \left(t = T^-_{(\text{obs})}\right)$

Setting 2

$H_0 : F(x) \geq F^*(x)$ for all $x \in (-\infty, \infty)$

$H_a : F(x) < F^*(x)$ for at least one value of x

Test statistic: $T^+ = \sup_x \left(F^*(x) - S(x) \right)$

Reject H_0 if $T^+_{(\text{obs})} > w_{1-\alpha}$. (Table A13, one-sided test)

p-value $= P\left(T^+ \geq T^+_{(\text{obs})} \right)$

Exact p-value $= t \sum_{j=0}^{[n(1-t)]} \binom{n}{j} \left(1 - t - \frac{j}{n} \right)^{n-j} \left(t + \frac{j}{n} \right)^{j-1} \quad \left(t = T^+_{(\text{obs})} \right)$

Setting 3

$H_0 : F(x) = F^*(x)$ for all $x \in (-\infty, \infty)$

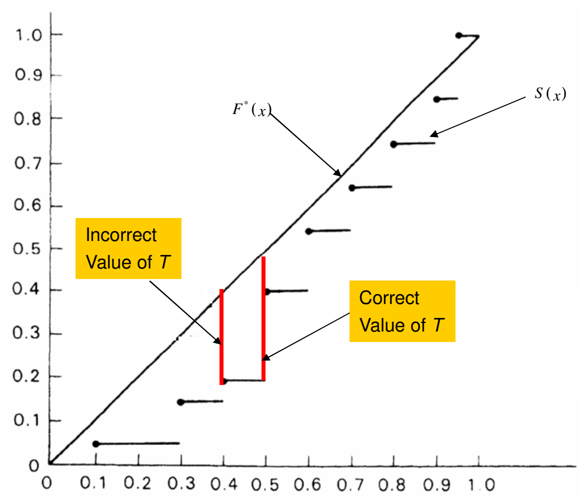
$H_a : F(x) \neq F^*(x)$ for at least one value of x

Test statistic: $T = \sup_x \left| S(x) - F^*(x) \right|$

Reject H_0 if $T_{(\text{obs})} > w_{1-\alpha}$. (Table A13, two-sided test)

p-value $= P\left(T \geq T_{(\text{obs})} \right)$

Exact p-value $= 2\, t \sum_{j=0}^{[n(1-t)]} \binom{n}{j} \left(1 - t - \frac{j}{n} \right)^{n-j} \left(t + \frac{j}{n} \right)^{j-1} \quad \left(t = T_{(\text{obs})} \right)$



Test statistic: $T = \sup_x \left| S(x) - F^*(x) \right|$

$$T = \max_{1 \le i \le n} \left\{ \max \left[\left| S(x_i) - F^*(x_i) \right|, \left| S(x_{i-1}) - F^*(x_i) \right| \right] \right\}$$

EXAMPLE 1

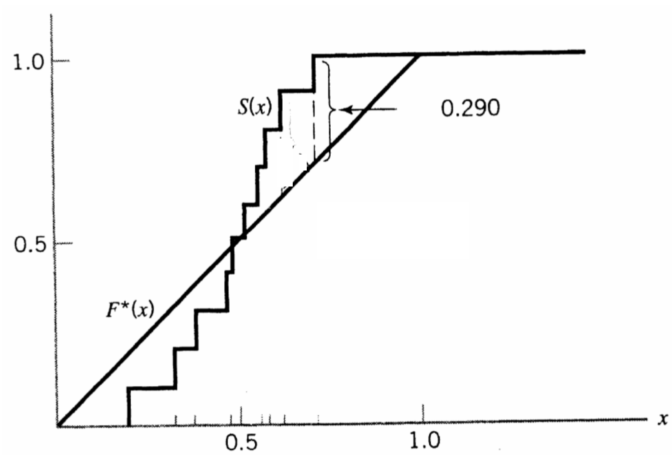
A random sample of size 10 is obtained: $X_1 = 0.621$, $X_2 = 0.503$, $X_3 = 0.203$, $X_4 = 0.477$, $X_5 = 0.710$, $X_6 = 0.581$, $X_7 = 0.329$, $X_8 = 0.480$, $X_9 = 0.554$, $X_{10} = 0.382$. The null hypothesis is that the distribution function is the uniform distribution function. The mathematical expression for the hypothesized distribution function is

$$F^*(x) = \begin{cases} 0 & x < 0 \\ x & 0 \le x < 1 \\ 1 & x \ge 1. \end{cases}$$

$H_0 : F(x) = F^*(x)$ for all $x \in (-\infty, \infty)$

$H_a : F(x) \ne F^*(x)$ for at least one value of x

$$T = \sup_x \left| S(x) - F^*(x) \right|$$



x_i	$S(x_i)$	$F^*(x_i)$	$\left\lvert S(x_i)-F^*(x_i)\right\rvert$	$\left\lvert S(x_{i-1})-F^*(x_i)\right\rvert$
0.203	0.1	0.203	0.103	–
0.329	0.2	0.329	0.129	0.229
0.382	0.3	0.382	0.082	0.182
0.477	0.4	0.477	0.077	0.177
0.480	0.5	0.480	0.020	0.080
0.503	0.6	0.503	0.097	0.003
0.554	0.7	0.554	0.146	0.046
0.581	0.8	0.581	0.219	0.119
0.621	0.9	0.621	0.279	0.179
0.710	1.0	0.710	0.290	0.190

- Confidence Band for the Population Distribution Function

Step 1. Draw a graph of the empirical distribution function $S(x)$ based on the random sample.

Step 2. Find the $1-\alpha$ quantile, $w_{1-\alpha}$, of the Kolmogorov test statistic from Table A13 for the two-sided test and for the appropriate sample size n.

Step 3. Draw a graph $(U(x))$ which is of a distance $w_{1-\alpha}$ above $S(x)$. This is the upper boundary of a $1-\alpha$ confidence band of $F(x)$.

Note: The value of $U(x)$ should not be over 1.

$$U(x)=\begin{cases} S(x)+w_{1-\alpha} & S(x)+w_{1-\alpha}\le 1 \\ 1 & S(x)+w_{1-\alpha}>1 \end{cases}$$

Step 4. Draw a graph $(V(x))$ which is of a distance $w_{1-\alpha}$ below $S(x)$. This is the lower boundary of a $1-\alpha$ confidence band of $F(x)$.

Note: The value of $V(x)$ should not be below 0.

$$V(x)=\begin{cases} S(x)-w_{1-\alpha} & S(x)-w_{1-\alpha}\ge 0 \\ 0 & S(x)-w_{1-\alpha}<0 \end{cases}$$

EXAMPLE 3

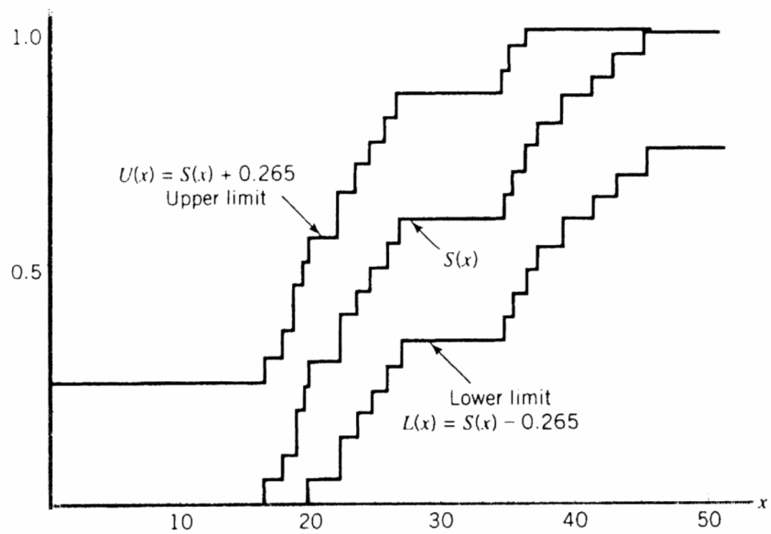
Suppose we wish to form a 90% confidence band for an unknown distribution function $F(x)$. A random sample of size 20 is obtained from the population with that distribution function. The results are ordered from smallest to largest for convenience.

16.7	17.4	18.1	18.2	18.8	19.3	22.4	22.4	24.0	24.7
25.7	27.0	35.1	35.8	36.5	37.6	39.8	42.1	43.2	46.2

$w_{1-\alpha} = w_{0.9} = 0.265$ (Table A13, two-sided)

x	16.7	17.4	18.1	18.2	18.8	19.3	22.4	22.4	24.0	24.7
$S(x)$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
$S(x)+w_{1-\alpha}$	0.315	0.365	0.415	0.465	0.515	0.565	0.615	0.665	0.715	0.765
$S(x)-w_{1-\alpha}$	–	–	–	–	–	0.035	0.085	0.135	0.185	0.235

x	25.7	27.0	35.1	35.8	36.5	37.6	39.8	42.1	43.2	46.2
$S(x)$	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.00
$S(x)+w_{1-\alpha}$	0.815	0.865	0.915	0.965	–	–	–	–	–	–
$S(x)-w_{1-\alpha}$	0.285	0.335	0.385	0.435	0.485	0.535	0.585	0.635	0.685	0.735



Section 6.2 Goodness-of-fit tests for families of distributions

When the Kolmogorov goodness-of-fit test is used, it is required that the hypothesized distribution is completely known. The goodness-of-fit tests introduced in this section allow the hypothesized distribution to have unknown parameters.

- Lilliefors Test for Normality

Let $X_1, X_2, \cdots, X_n$ be a random sample from a population distribution.

H_0: The population distribution is a normal distribution with unknown mean μ and standard deviation σ. $\left(X \sim N\left(\mu, \sigma^2\right)\right)$

H_a: The population distribution is not normal.

Estimate of μ: $\hat{\mu} = \bar{x} = \dfrac{\sum_{i=1}^{n} x_i}{n}$

Estimate of σ: $\hat{\sigma} = s = \sqrt{\dfrac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n-1}}$

Normalized sample values: $z_i = \dfrac{x_i - \bar{x}}{s}$ $\left(i = 1, 2, \cdots, n\right)$

Let $S(x)$ be the empirical distribution function based on $z_1, z_2, \cdots, z_n$.

Test statistic: $T_1 = \sup_x \left| F^*(x) - S(x) \right|$

$\left(F^*(x)\right.$ is the distribution function of a standard normal distribution.$\left.\right)$

Reject H_0 if $T_{1(\mathrm{obs})} > w_{1-\alpha}$. (Table A14)

p-value $= P\left(T_1 \geq T_{1(\mathrm{obs})}\right)$

EXAMPLE 1 (Data used in Example 4.5.3)

Fifty two-digit numbers were drawn at random from a telephone book, and the chi-squared test for goodness of fit is used to see if they could have been observations on a normally distributed random variable. The numbers, after being arranged in order from the smallest to the largest, are as follows.

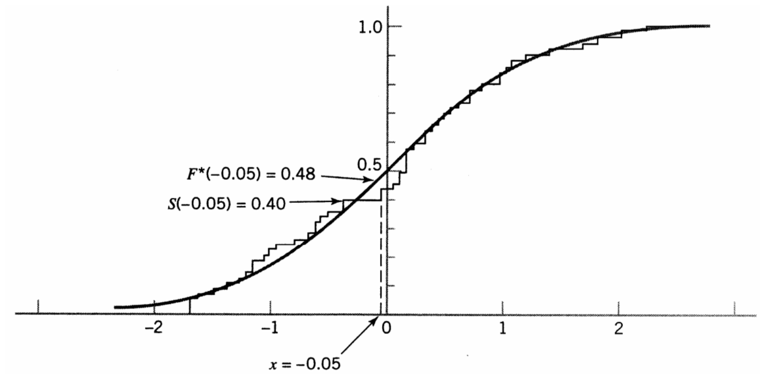
23 23 24 27 29 31 32 33 33 35
36 37 40 42 43 43 44 45 48 48
54 54 56 57 57 58 58 58 58 59
61 61 62 63 64 65 66 68 68 70
73 73 74 75 77 81 87 89 93 97

H_0 : These numbers are from a normal distribution.

H_a : These numbers are not from a normal distribution.

x_i	z_i	x_i	z_i	x_i	z_i	x_i	z_i	x_i	z_i
23	−1.69	36	−1.00	54	−0.05	61	0.31	73	0.95
23	−1.69	37	−0.95	54	−0.05	61	0.31	73	0.95
24	−1.63	40	−0.79	56	0.05	62	0.37	74	1.00
27	−1.48	42	−0.69	57	0.10	63	0.42	75	1.05
29	−1.37	43	−0.63	57	0.10	64	0.47	77	1.16
31	−1.27	43	−0.63	58	0.16	65	0.52	81	1.37
32	−1.21	44	−0.58	58	0.16	66	0.58	87	1.68
33	−1.16	45	−0.53	58	0.16	68	0.68	89	1.79
33	−1.16	48	−0.37	58	0.16	68	0.68	93	2.00
35	−1.05	48	−0.37	59	0.21	70	0.79	97	2.21

Let $S(x)$ be the empirical distribution function based on $z_1, z_2, \cdots, z_n$.



$T_1 = \sup_x \left| F^*(x) - S(x) \right| = 0.08$

- Lilliefors Test for the Exponential Distribution

  H_0: The population distribution is an exponential distribution with unknown mean θ. $\left(F(x) = 1 - \exp\left(-\frac{x}{\theta}\right) \; (x > 0) \right)$

  H_a: The population distribution is not exponential.

  Estimate of θ: $\hat{\theta} = \overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$

  $z_i = \frac{x_i}{\overline{x}} \quad (i = 1, 2, \cdots, n)$

  Let $S(x)$ be the empirical distribution function based on $z_1, z_2, \cdots, z_n$.

Test statistic: $T_2 = \sup_x \left| F^*(x) - S(x) \right|$

$\left(F^*(x) = 1 - \exp(-x) \;\; (x > 0)\right)$

Reject H_0 if $T_{2\,(\text{obs})} > w_{1-\alpha}$. (Table A15)

p-value $= P\left(T_2 \ge T_{2\,(\text{obs})}\right)$

EXAMPLE 2

The placement of long-distance telephone calls through a certain switchboard is believed to be a random process, with times between calls having an exponential distribution. The first 10 calls after 1 P.M. one Monday occurred at 1:06, 1:08, 1:16, 1:22, 1:23, 1:34, 1:44, 1:47, 1:51, and 1:57. The successive times between calls, counting the first time from l:00 to 1:06, are (in minutes):

6, 2, 8, 6, 1, 11, 10, 3, 4, 6.

Test the hypothesis that the data were from an exponential distribution using $\alpha = 0.05$.

i	x_i	$z_i = x_i / \bar{x}$	$1-\exp(-z_i)$	$i/10-(1-\exp(-z_i))$	$(1-\exp(-z_i))-(i-1)/10$
1	1	0.1754	0.1609	−0.0609	0.1609
2	2	0.3508	0.2959	−0.0959	0.1959
3	3	0.5263	0.4092	−0.1092	0.2092
4	4	0.7018	0.5043	−0.1043	0.2043
5	6	1.0526	0.6510	−0.1510	0.2510
6	6	1.0526	0.6510	−0.0510	0.1510
7	6	1.0526	0.6510	0.0490	0.0510
8	8	1.4035	0.7543	0.0457	0.0543
9	10	1.7544	0.8270	0.0730	0.0270
10	11	1.9298	0.8548	0.1452	−0.0452

- Shapiro-Wilk Test for Normality

  Let $X_1, X_2, \cdots, X_n$ be a random sample from a population distribution.

  H_0: The population distribution is a normal distribution with unknown mean μ and standard deviation σ. $\left(X \sim N(\mu, \sigma^2)\right)$

  H_a: The population distribution is not normal.

  Steps for calculating W statistic T_3:

  Step 1. Calculate $D = \sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2$.

  Step 2. Find coefficients $a_1, a_2, \cdots, a_k$ from Table A16. $\left(k = [n/2]\right)$

  Step 3. Find the value of the test statistic: $T_3 = W = \dfrac{\left[\sum_{i=1}^{k} a_i \left(X^{(n-i+1)} - X^{(i)}\right)\right]^2}{D}$.

  $\left(X^{(1)}, X^{(2)}, \cdots, X^{(n)}\right.$ are the order statistics.$\left.\right)$

Reject H_0 if $T_{3(\text{obs})} = W_{(\text{obs})} < w_\alpha$. (Table A17, $n \le 50$)

Approximate p-value $= P(Z \le G)$

Case A. $3 \le n \le 6$

Calculate $v = \ln\left(\frac{W - d_n}{1 - W}\right)$.

(The value of d_n can be found from Table A18.)

For calculated value of v, find the value of G using the same table.

Case B. $7 \le n \le 50$

$$G = b_n + c_n \ln\left(\frac{W - d_n}{1 - W}\right)$$

(The values of b_n, c_n, d_n can be found from Table A18.)

EXAMPLE 3 (Data used in Example 4.5.3 and Example 6.2.1)

Fifty two-digit numbers were drawn at random from a telephone book, and the chi-squared test for goodness of fit is used to see if they could have been observations on a normally distributed random variable. The numbers, after being arranged in order from the smallest to the largest, are as follows.

23	23	24	27	29	31	32	33	33	35
36	37	40	42	43	43	44	45	48	48
54	54	56	57	57	58	58	58	58	59
61	61	62	63	64	65	66	68	68	70
73	73	74	75	77	81	87	89	93	97

H_0 : These numbers are from a normal distribution.

H_a : These numbers are not from a normal distribution.

i	a_i	$X^{(n-i+1)} - X^{(i)}$
1	0.3751	97−23
2	0.2574	93−23
3	0.2260	89−24
4	0.2032	87−27
5	0.1847	81−29
6	0.1691	77−31
7	0.1554	75−32
8	0.1430	74−33
9	0.1317	73−33
10	0.1212	73−35
11	0.1113	70−36
12	0.1020	68−37
13	0.0932	68−40

i	a_i	$X^{(n-i+1)} - X^{(i)}$
14	0.0846	66−42
15	0.0764	65−43
16	0.0685	64−43
17	0.0608	63−44
18	0.0532	62−45
19	0.0459	61−48
20	0.0386	61−48
21	0.0314	59−54
22	0.0244	58−54
23	0.0174	58−56
24	0.0104	58−57
25	0.0035	58−57

- Combining Several Independent Test Results

  Let $T_1, T_2, \cdots, T_m$ $(\text{or } W_1, W_2, \cdots, W_m)$ be the values of m independent Shapiro-Wilk test results for normality (or log normality). It is desired to combine the test results.

  Step 1. Convert $T_1, T_2, \cdots, T_m$ $(\text{or } W_1, W_2, \cdots, W_m)$ to $G_1, G_2, \cdots, G_m$ using Table A18.

  Step 2. Find the p-value of the combined test. $p\text{-value} = P\left(Z \le \frac{\sum_{i=1}^{m} G_i}{\sqrt{m}} \right)$

EXAMPLE 4

When an offshore lease is made available for bids, several oil companies usually submit bids for the right to drill for oil in that area. The distribution of these bids is often assumed to follow the "lognormal" distribution; that is, the logarithm of the bids is assumed to follow the normal distribution. However, the means and variances may vary from lease to lease. Also, the number of bids on any one lease is usually too small to be able to tell whether the normality assumption on the logarithms of the bids is reasonable or not. The purpose is to test the hypothesis:

H_0: The bids are lognormally distributed,

H_a: The bids are not lognormally distributed.

Lease Number	Number of Bids	$T_3(W)$	G
1	14	0.9243	−0.6550
2	14	0.9757	1.3559
3	14	0.9717	1.0939
4	14	0.8772	−1.5848
5	14	0.9537	0.2345
6	15	0.9135	−1.0093
7	15	0.8629	−1.9321
8	15	0.8786	−1.6806
9	15	0.8515	−2.1011
10	15	0.9226	−0.7966
11	15	0.9581	0.3354
12	15	0.9625	0.5344
13	16	0.9178	−1.0151
14	16	0.8596	−2.1011
15	15	0.9603	0.4323
16	16	0.9669	0.6795
		Total	−8.2099

Section 6.3 Tests on two independent samples

- Smirnov Test
  (Kolmogorov-Smirnov Two-Sample Test)

  Let $x_1, x_2, \cdots, x_n$ be the observations of a random sample from a population with distribution function $F(.)$.

  Let $y_1, y_2, \cdots, y_m$ be the observations of a random sample from a population with distribution function $G(.)$.

  It is assumed that the two samples are independent.

Setting 1

$H_0: F(x) \le G(x)$ for all $x \in (-\infty, \infty)$

$H_a: F(x) > G(x)$ for at least one value of x

(X values tend to be smaller than the Y values.)

Setting 2

$H_0: F(x) \ge G(x)$ for all $x \in (-\infty, \infty)$

$H_a: F(x) < G(x)$ for at least one value of x

(X values tend to be larger than the Y values.)

Setting 3

$H_0: F(x) = G(x)$ for all $x \in (-\infty, \infty)$

$H_a: F(x) \ne G(x)$ for at least one value of x

Setting 1

$H_0: F(x) \le G(x)$ for all $x \in (-\infty, \infty)$

$H_a: F(x) > G(x)$ for at least one value of x

Test statistic: $T_1^+ = \sup_x \left(S_1(x) - S_2(x)\right)$

Reject H_0 if $T_{1(\mathrm{obs})}^+ > w_{1-\alpha}$.

(Table A19 $(n = m)$ or Table A20 $(n \ne m)$, one-sided test)

p-value $= P\left(T_1^+ \ge T_{1(\mathrm{obs})}^+\right)$

Exact p-value when $n = m$:

$$p\text{-value} = \frac{\binom{2n}{n + \left[n\, T_{1(\mathrm{obs})}^+\right]}}{\binom{2n}{n}}$$

Setting 2

$H_0: F(x) \geq G(x)$ for all $x \in (-\infty, \infty)$

$H_a: F(x) < G(x)$ for at least one value of x

Test statistic: $T_1^- = \sup_x \left(S_2(x) - S_1(x) \right)$

Reject H_0 if $T_{1(\text{obs})}^- > w_{1-\alpha}$.

(Table A19 $(n = m)$ or Table A20 $(n \neq m)$, one-sided test)

p-value $= P\left(T_1^- \geq T_{1(\text{obs})}^-\right)$

Exact p-value when $n = m$:

$$p\text{-value} = \frac{\binom{2n}{n + \left[n\, T_{1(\text{obs})}^- \right]}}{\binom{2n}{n}}$$

Setting 3

$H_0: F(x) = G(x)$ for all $x \in (-\infty, \infty)$

$H_a: F(x) \neq G(x)$ for at least one value of x

Test statistic: $T_1 = \sup_x \left| S_1(x) - S_2(x) \right|$

Reject H_0 if $T_{1(\text{obs})} > w_{1-\alpha}$.

(Table A19 $(n = m)$ or Table A20 $(n \neq m)$, two-sided test)

p-value $= P\left(T_1 \geq T_{1(\text{obs})}\right)$

Exact p-value when $n = m$:

$$p\text{-value} = \frac{2\binom{2n}{n + \left[n\, T_{1(\text{obs})} \right]}}{\binom{2n}{n}}$$

EXAMPLE 1

A random sample $X_1, X_2, \cdots, X_9$ is obtained from one population, and a random sample $Y_1, Y_2, \cdots, Y_{15}$ is obtained from a second population. The ordered X and Y values are listed here.

X	Y
7.6	5.2
8.4	5.7
8.6	5.9
8.7	6.5
9.3	6.8
9.9	8.2
10.1	9.1
10.6	9.8
11.2	10.8
	11.3
	11.5
	12.3
	12.5
	13.4
	14.6

X_i	Y_i	$S_1(x)-S_2(x)$
	5.2	$0-1/15=-1/15$
	5.7	$0-2/15=-2/15$
	5.9	$0-3/15=-1/5$
	6.5	$0-4/15=-4/15$
	6.8	$0-5/15=-1/3$
7.6		$1/9-5/15=-2/9$
	8.2	$1/9-6/15=-13/45$
8.4		$2/9-6/15=-8/45$
8.6		$3/9-6/15=-1/15$
8.7		$4/9-6/15=2/45$
	9.1	$4/9-7/15=-1/45$
9.3		$5/9-7/15=4/45$
	9.8	$5/9-8/15=1/45$
9.9		$6/9-8/15=2/15$
10.1		$7/9-8/15=11/45$
10.6		$8/9-8/15=16/45$
	10.8	$8/9-9/15=13/45$
11.2		$1-9/15=2/5$
	11.3	$1-10/15=1/3$
	11.5	$1-11/15=4/15$
	12.3	$1-12/15=1/5$
	12.5	$1-13/15=2/15$
	13.4	$1-14/15=1/15$
	14.6	$1-1=0$

- Cramer-von Mises Two-Sample Test

Let $x_1, x_2, \cdots, x_n$ be the observations of a random sample from a population with distribution function $F(\cdot)$.

Let $y_1, y_2, \cdots, y_m$ be the observations of a random sample from a population with distribution function $G(\cdot)$.

It is assumed that the two samples are independent.

$H_0 : F(x) = G(x)$ for all $x \in (-\infty, \infty)$

$H_a : F(x) \neq G(x)$ for at least one value of x

Test statistic

$$T_2 = \frac{mn}{(m+n)^2}\left\{\sum_{i=1}^{n}\left[S_1(X_i) - S_2(X_i)\right]^2 + \sum_{i=1}^{m}\left[S_1(Y_i) - S_2(Y_i)\right]^2\right\}$$

Reject H_0 if $T_{2(\text{obs})} > w_{1-\alpha}$.

$w_{0.10} = 0.046$	$w_{0.50} = 0.119$	$w_{0.90} = 0.347$
$w_{0.20} = 0.062$	$w_{0.60} = 0.147$	$w_{0.95} = 0.461$
$w_{0.30} = 0.079$	$w_{0.70} = 0.184$	$w_{0.99} = 0.743$
$w_{0.40} = 0.097$	$w_{0.80} = 0.241$	$w_{0.999} = 1.168$

EXAMPLE 2 (Data used in Example 1)

A random sample of $X_1, X_2, \cdots, X_9$ is obtained from one population, and a random sample of $Y_1, Y_2, \cdots, Y_{15}$ is obtained from a second population. The ordered X and Y values are listed here.

$H_0 : F(x) = G(x)$ for all $x \in (-\infty, \infty)$

$H_a : F(x) \neq G(x)$ for at least one value of x

$$T_2 = \frac{mn}{(m+n)^2}\left\{\sum_{i=1}^{n}\left[S_1(X_i) - S_2(X_i)\right]^2 + \sum_{i=1}^{m}\left[S_1(Y_i) - S_2(Y_i)\right]^2\right\}$$

X	Y
7.6	5.2
8.4	5.7
8.6	5.9
8.7	6.5
9.3	6.8
9.9	8.2
10.1	9.1
10.6	9.8
11.2	10.8
	11.3
	11.5
	12.3
	12.5
	13.4
	14.6

X_i	Y_i	$S_1(x) - S_2(x)$
	5.2	0−1/15 = −1/15
	5.7	0−2/15 = −2/15
	5.9	0−3/15 = −1/5
	6.5	0−4/15 = −4/15
	6.8	0−5/15 = −1/3
7.6		1/9−5/15 = −2/9
	8.2	1/9− 6/15 = −13/45
8.4		2/9−6/15 = −8/45
8.6		3/9−6/15 = −1/15
8.7		4/9−6/15 = 2/45
	9.1	4/9−7/15 = −1/45
9.3		5/9−7/15 = 4/45
	9.8	5/9−8/15 = 1/45
9.9		6/9−8/15 = 2/15
10.1		7/9−8/15 = 11/45
10.6		8/9−8/15 = 16/45
	10.8	8/9−9/15 = 13/45
11.2		1−9/15 = 2/5
	11.3	1−10/15 = 1/3
	11.5	1−11/15 = 4/15
	12.3	1−12/15 = 1/5
	12.5	1−13/15 = 2/15
	13.4	1−14/15 = 1/15
	14.6	1−1 = 0

$$\sum_{i=1}^{n}\left[S_1(X_i) - S_2(X_i)\right]^2 = 0.459$$

X_i	Y_i	$S_1(x)-S_2(x)$
	5.2	0−1/15 = −1/15
	5.7	0−2/15 = −2/15
	5.9	0−3/15 = −1/5
	6.5	0−4/15 = −4/15
	6.8	0−5/15 = −1/3
7.6		1/9−5/15 = −2/9
	8.2	1/9− 6/15 = −13/45
8.4		2/9−6/15 = −8/45
8.6		3/9−6/15 = −1/15
8.7		4/9−6/15 = 2/45
	9.1	4/9−7/15 = −1/45
9.3		5/9−7/15 = 4/45
	9.8	5/9−8/15 = 1/45
9.9		6/9−8/15 = 2/15
10.1		7/9−8/15 = 11/45
10.6		8/9−8/15 = 16/45
	10.8	8/9−9/15 = 13/45
11.2		1−9/15 = 2/5
	11.3	1−10/15 = 1/3
	11.5	1−11/15 = 4/15
	12.3	1−12/15 = 1/5
	12.5	1−13/15 = 2/15
	13.4	1−14/15 = 1/15
	14.6	1−1 = 0

$w_{0.10}=0.046$	$w_{0.50}=0.119$	$w_{0.90}=0.347$
$w_{0.20}=0.062$	$w_{0.60}=0.147$	$w_{0.95}=0.461$
$w_{0.30}=0.079$	$w_{0.70}=0.184$	$w_{0.99}=0.743$
$w_{0.40}=0.097$	$w_{0.80}=0.241$	$w_{0.999}=1.168$