Random Numbers Testing

Recall

- The main two properties of random numbers are uniformity and independence.
- Several tests help in validating if these properties are achieved or not

1. Frequency test: Uses Kolmogorov-Smirnov or Chi-square tests to compare the distribution of PRN to the uniform distribution
2. Autocorrelation test: Test the correlation between numbers and compare it to the expected correlation

Hypotheses

- In testing the uniformity, there are two hypotheses:

$$H_0: R \sim Uniform[0-1]$$
$$H_1: R \nsim Uniform[0-1]$$

- H_0 is called the null hypothesis, and it means that:
  - The generated random numbers follow the uniform distribution
  - There is **no significant** difference between the distribution of the RN and the Uniform distribution
- Sometimes we need further tests to ensure the uniformity

Hypotheses

- In testing the independence, there are two hypotheses:

$$H_0: R \sim independency$$
$$H_1: R \nsim independency$$

The significance level α

- This is an important parameter for any statistical test, and it must be defined before starting the test
- It refers to the probability of rejecting the null hypothesis, given that it is true

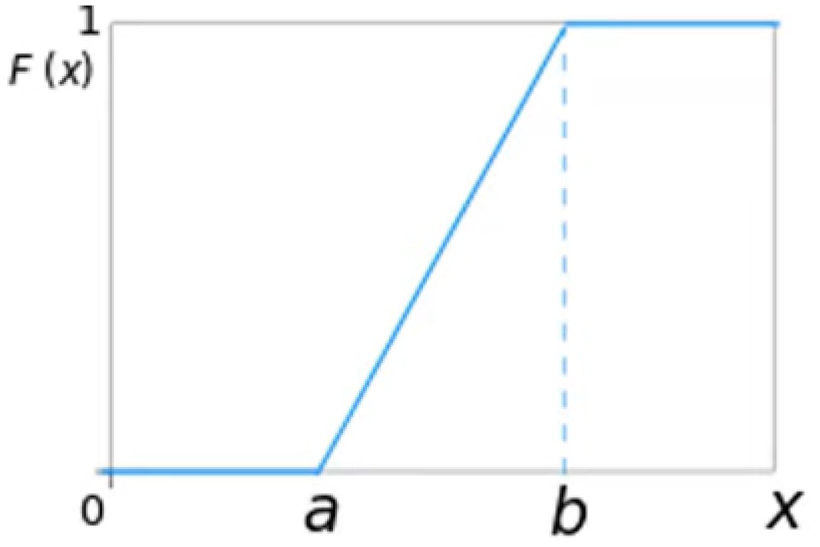
$$\alpha = P(reject H_0 | H_0\ true)$$

- α is set by the decision makers

K-S test

- Compares the cumulative distribution function of the uniform distribution $F(x)$, to the distribution of the generated sample, $S_N(x)$ with Nobservations, of random numbers

$$F(x) = \begin{cases} 0, & x \leq a \\ \dfrac{x-a}{b-a}, & a < x < b \\ 1, & x \geq b \end{cases}$$



Cont.

- Compares the cumulative distribution function of the uniform distribution $F(x)$, to the distribution of the generated sample, $S_N(x)$ with Nobservations, of random numbers

$$D = \max(|F(x) - S_N(x)|)$$

- The sampling distribution of D is known, and defined as a function of N

Cont.

Kolmogorov--Smirnov Critical Values

Degrees of Freedom (N)	$D_{0.10}$	$D_{0.05}$	$D_{0.01}$
1	0.950	0.975	0.995
2	0.776	0.842	0.929
3	0.642	0.708	0.828
4	0.564	0.624	0.733
5	0.510	0.565	0.669
6	0.470	0.521	0.618
7	0.438	0.486	0.577
8	0.411	0.457	0.543
9	0.388	0.432	0.514
10	0.368	0.410	0.490
11	0.352	0.391	0.468
12	0.338	0.375	0.450
13	0.325	0.361	0.433
14	0.314	0.349	0.418
15	0.304	0.338	0.404
16	0.295	0.328	0.392
17	0.286	0.318	0.381
18	0.278	0.309	0.371
19	0.272	0.301	0.363
20	0.264	0.294	0.356
25	0.24	0.27	0.32
30	0.22	0.24	0.29
35	0.21	0.23	0.27
Over 35	$\frac{1.22}{\sqrt{N}}$	$\frac{1.36}{\sqrt{N}}$	$\frac{1.63}{\sqrt{N}}$

Cont.

- The critical value is extracted from the table based on N and $alpha$

Kolmogorov--Smirnov Critical Values

$if\ N = 8$ and $\alpha = 0.05$
$D_{0.05} = 0.457$

Degrees of Freedom (*N*)	$D_{0.10}$	$D_{0.05}$	$D_{0.01}$
1	0.950	0.975	0.995
2	0.776	0.842	0.929
3	0.642	0.708	0.828
4	0.564	0.624	0.733
5	0.510	0.565	0.669
6	0.470	0.521	0.618
7	0.438	0.486	0.577
8	0.411	0.457	0.543
9	0.388	0.432	0.514
10	0.368	0.410	0.490
[illegible]	[illegible]	.391	0.468
[illegible]	[illegible]	.375	0.450
13	0.325	0.361	0.433
14	0.314	0.349	0.418
15	0.304	0.338	0.404
16	0.295	0.328	0.392
17	0.286	0.318	0.381
18	0.278	0.309	0.371
19	0.272	0.301	0.363
20	0.264	0.294	0.356
25	0.24	0.27	0.32
30	0.22	0.24	0.29
35	0.21	0.23	0.27
Over 35	$\frac{1.22}{\sqrt{N}}$	$\frac{1.36}{\sqrt{N}}$	$\frac{1.63}{\sqrt{N}}$

Calculation steps

- The following steps lead to validate the uniformity of the generated RN, according to the K-S test:

1. Sort the random number in an ascending order
2. Define the null hypothesis $H_0: R \sim U[0-1]$
3. Calculate D^+ and D^- as follows:

$$D^+ = max\left\{\frac{i}{N} - R_i\right\}, \text{ and } D^- = max\left\{R_i - \frac{i-1}{N}\right\}$$

4. Calculate D as $D = max\{D^+, D^-\}$
5. Locate the critical value, based on N and α
6. Compare D to the critical value D_α

Cont.

- Finally, If the sample statistic D **is greater than** the critical value D_α, the null hypothesis is rejected.
- If $D \ < D_\alpha$, conclude that we fail to reject the null hypothesis and these PRNs can be accepted according to K-S test.

Example

- Assume that we have generated 5 random numbers: $0.44, 0.81, 0.14, 0.05, 0.93$ and we want to test their uniformity under the level of significance 0.05

- We know from the above example that N is 5 and α is 0.05

Extract the critical

$N = 5$ and $\alpha = 0.05$
$D_{0.05} = 0.565$

Degrees of Freedom (*N*)	$D_{0.10}$	$D_{0.05}$	$D_{0.01}$
1	0.950	0.975	0.995
2	0.776	0.842	0.929
3	0.642	0.708	0.828
4	0.564	0.624	0.733
5	0.510	0.565	0.669
6	0.470	0.521	0.618
7	0.438	0.486	0.577
8	0.411	0.457	0.543
9	0.388	0.432	0.514
10	0.368	0.410	0.490

Calculate D

- The first step was to sort the random numbers, i is values from 1 to N

R_i	0.05	0.14	0.44	0.81	0.93

Cont.

R_i	0.05	0.14	0.44	0.81	0.93
i/N					

Cont.

R_i	0.05	0.14	0.44	0.81	0.93
i/N	0.20	0.40	0.60	0.80	1.00
$i/N - R_i$					

Cont.

R_i	0.05	0.14	0.44	0.81	0.93
i/N	0.20	0.40	0.60	0.80	1.00
$i/N - R_i$	0.15	**0.26**	0.16	–	0.07
$(i-1)/N$					

Cont.

$$D^+ = max\left\{\frac{i}{N} - R_i\right\}, \text{ and } D^- = max\left\{R_i - \frac{i-1}{N}\right\}$$

R_i	0.05	0.14	0.44	0.81	0.93
i/N	0.20	0.40	0.60	0.80	1.00
$i/N - R_i$	0.15	**0.26**	0.16	–	0.07
$(i-1)/N$	0.00	0.20	0.40	0.60	0.80
$R_i - (i-1)/N$					

Decision

Since the computed value, 0.26, is less than the critical value, 0.565, the hypothesis that the distribution of the generated numbers is the uniform distribution is "***failed to reject***".

$$D = \max\{D^+, D^-\} = 0.26 < D_\alpha$$

$$\boxed{D_\alpha = 0.565}$$

x^2 test

- For validating the RN using this method, we follow the next steps:

1. divide the interval [0-1] into k intervals, typically, this should be 100 intervals at least
2. formulate the null hypothesis, which assumes the uniformity
3. calculate x^2 value:

$$x^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

Cont.

is the observed #
in the i^{th} class

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

is the expected #
in the i^{th} class

How many $R's$
in the i^{th} class

k is the #
of classes

$$E_i = \frac{N}{k}$$

Uniform
Distribution

Cont.

4. Extract the x_α^2 critical value from the x_α^2 distribution table

this done using $k - 1$ degree of freedom and significance level α

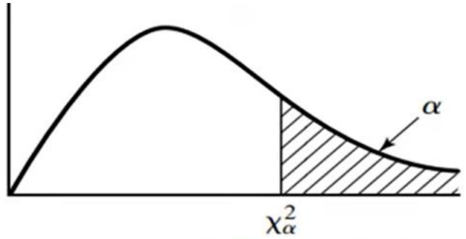
This value is denoted by $x_{\alpha,\,k-1}^2$

5. Compare x^2 to $x_{\alpha,\,k-1}^2$

if $x^2 > x_{\alpha,\,k-1}^2$, we reject the null hypothesis, otherwise, we fail to reject

x^2 table

$\chi^2_{0.025,\,9} = 19$



$\alpha = 0.025,$
$k = 10$

v	$\chi^2_{0.005}$	$\chi^2_{0.01}$	$\chi^2_{0.025}$	$\chi^2_{0.05}$	$\chi^2_{0.10}$
1	7.88	6.63	5.02	3.84	2.71
2	10.60	9.21	7.38	5.99	4.61
3	12.84	11.34	9.35	7.81	6.25
4	14.96	13.28	11.14	9.49	7.78
5	16.7	15.1	12.8	11.1	9.2
6	18.5	16.8	14.4	12.6	10.6
7	20.3	18.5	16.0	14.1	12.0
8	22.0	20.1	17.5	15.5	13.4
9	23.6	21.7	19.0	16.9	14.7
10	25.2	23.2	20.5	18.3	16.0
11	26.8	24.7	21.9	19.7	17.3
12	28.3	26.2	23.3	21.0	18.5
13	29.8	27.7	24.7	22.4	19.8
14	31.3	29.1	26.1	23.7	21.1
15	32.8	30.6	27.5	25.0	22.3
16	34.3	32.0	28.8	26.3	23.5
17	35.7	33.4	30.2	27.6	24.8
18	37.2	34.8	31.5	28.9	26.0
19	38.6	36.2	32.9	30.1	27.2

Example

We have 100 generated PRNs R_i's are shown below. Use Chi-square test with $\alpha = 0.05$ and $k = 10$.

0.34	0.90	0.25	0.89	0.87	0.44	0.12	0.21	0.46	0.67
0.83	0.76	0.79	0.64	0.70	0.81	0.94	0.74	0.22	0.74
0.96	0.99	0.77	0.67	0.56	0.41	0.52	0.73	0.99	0.02
0.47	0.30	0.17	0.82	0.56	0.05	0.45	0.31	0.78	0.05
0.79	0.71	0.23	0.19	0.82	0.93	0.65	0.37	0.39	0.42
0.99	0.17	0.99	0.46	0.05	0.66	0.10	0.42	0.18	0.49
0.37	0.51	0.54	0.01	0.81	0.28	0.69	0.34	0.75	0.49
0.72	0.43	0.56	0.97	0.30	0.94	0.96	0.58	0.73	0.05
0.06	0.39	0.84	0.24	0.40	0.64	0.40	0.19	0.79	0.62
0.18	0.26	0.97	0.88	0.64	0.47	0.60	0.11	0.29	0.78

Cont.

	Interval	O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
[0.0, 0.1)	1	8	10	−2	4	0.4
[0.1, 0.2)	2	8	10	−2	4	0.4
[0.2, 0.3)	3	10	10	0	0	0.0
	4	9	10	−1	1	0.1
	5	12	10	2	4	0.4
	6	8	10	−2	4	0.4
	7	10	10	0	0	0.0
	8	14	10	4	16	1.6
	9	10	10	0	0	0.0
[0.9, 1.0)	10	11	10	1	1	0.1
		100	100	0		3.4

decision.

- $x^2 = 3.4$
- $x^2_{0.05,\, 9} = 16.9$

- $x^2 < x^2_{0.05,\, 9}$, therefore, we fail to reject the null.
  - The numbers belong to the uniform distribution

Autocorrelation tests

- Self-study for this semester