modelling simulation

Generation of random numbers → 1) random nos.
2) properties of R.N. < uniformity, independence.
3) True RN, Pseudo RN
4) ways to generate
5) characteristics
6) Errors.

used to study true behaviour of a system. eg: tossing coin, dice.

True random no.
↳ sources → physical process, existing table.

ways/Techniques to generate → LCM (Lehmer linear congruential method), LLCG, random no. stream

$c \neq 0$, mixed cm
$c = 0$, multiplicative cm

$$x_{i+1} = (a x_i + c) \bmod m, \quad i = 0, 1, 2 \dots$$

multiplier, increment, modulus.

convert int. to random nos. ⇒ $R_i = \frac{x_i}{m}$, $i = 1, 2 \dots$

mixed congruential method, generate a sequence of three 2 digit random numbers.

$x_0 = 27$, $a = 17$, $c = 43$, $m = 100$ $(i = 0)$

$x_1 = (27(17) + 43) \% 100$

$= (459 + 43) \% 100$

$= 502 \% 100$

$= 2$

$R_1 = \frac{x_1}{m} = \frac{2}{100} = 0.02$

$x_2 = (2(17) + 43) \% 100$

$= 77$

$R_2 = \frac{x_2}{m} = \frac{77}{100} = 0.77$

$x_3 = (77(17) + 43) \% 100$

$= 1352 \% 100 = 52$

$R_3 = 0.52$

$X_0 = 39,\ a = 7,\ c = 29,\ m = 100$

$X_1 = (39(7) + 29) \% 100$

$= 302 \% 100 = 2$

$R_1 = 0.02$

$X_2 = (\cancel{39}\,2 \times 7 + 29) \% 100$

$= 43 \% 100 = 43$

$R_2 = 0.43$

$X_3 = (43 \times 7 + 29) \% 100$

$= 330 \% 100 = 30$

$R_3 = 0.3$

characteristics

↳ max density

↳ max period

case 1 : $P = m = 2^b$

case 2 : $P = m/4 = 2^{b-2}$ [$\because c = 0$]

seed $X_0 \rightarrow$ odd

$a = 3 + 8k$ or $5 + 8k$.

for $k = 0, 1, 2$

$a = 13,\ m = 64,\ X_0 = 1$

$P = \frac{\cancel{64}\ 32}{\cancel{4}\ \cancel{2}} = 16$

~~(13 = 3 + 8(5k))~~ $X_0 = (aX_1 + c) \% m$

~~$1 = 13(X_1) \% 64$~~

$X_1 = 13(1) \% 64$

$= 13 \% 64$

$= 13$

$X_2 = 13(13) \% 64$

$= 169 \% 64$

$= 41$

MS.

$x_3 = 13(41) \% 64$
$= 533 \% 64$
$= 21$

$x_4 = 13(21) \% 64$
$= 273 \% 64$
$= 17$

$x_5 = 13(17) \% 64$
$= 221 \% 64$
$= 29$

$x_6 = 13(29) \% 64$
$= 377 \% 64 = 57$

$x_7 = 13(57) \% 64$
$= 741 \% 64 = 37$

$x_8 = 13(37) \% 64$
$= 481 \% 64 = 33$

$x_9 = 13(33) \% 64$
$= 429 \% 64 = 45$

$x_{10} = 13(45) \% 64$
$= 585 \% 64 = 9$

$x_{11} = 13(9) \% 64$
$= 53$

$x_{12} = 13(53) \% 64$
$= 49$

$x_{13} = 61$

$x_{14} = 25$

$x_{15} = 5$

$x_{16} = 1$

∴ it satisfies.

$a = 13,\ m = 64,\ x_0 = 2$

$x_1 = 13(2) \% 64 = 26$

$x_2 = 13(26) \% 64 = 18$

$x_3 = 13(18) \% 64 = 42$

$x_4 = 13(42) \% 64 = 34$

$x_5 = 13(34) \% 64 = 58$

$x_6 = 13(58) \% 64 = 50$

$x_7 = 13(50) \% 64 = 10$

$x_8 = 13(10) \% 64 = 2$

$x_9 = 13(2) \% 64 = 26$

$x_{10} = 13(26) \% 64 = 18$

$x_{11} = 42$

$x_{12} = 34$

$\vdots$

$x_{16} = 1$

$x_0 = 4$

$x_1 = 13(4) \% 64 = 52$

$x_2 = 13(52) \% 64 = 36$

$x_3 = 13(36) \% 64 = 20$

$x_4 = 13(20) \% 64 = 4.$

Case 3: m is prime, $c = 0$

longest possible period $P = m-1$, achieved whenever multiplier a has property that smallest int. k such that $a^k - 1$ is divisible by m is $k = m-1$

1 $m = 7$, $a = 3$, $x_0 = 5.$

$P = 6$

$x_1 = a(x_0) \% m$

$3(5) \% 7$

$15 \% 7 = 1$

$x_2 = 3(1) \% 7 = 3$

$x_3 = 3(3) \% 7 = 2$

$x_4 = 3(2) \% 7 = 6$

$x_5 = 3(6) \% 7 = 4$

$x_6 = 3(4) \% 7 = 5$

Introduction to simulation

↳ imitation of operatn of real world process/system over time

~~one checkout counter~~

customers arrives at random 1 to 8 min. Service time 1 to 6 min.

simulate for arrival of 20 customers. Analyze ~~for~~: based on:

1) Avg. waiting time 2) Prob. customer waits in queue.

Each possible value of inter arrival time ~~& service time~~ have same prob. of occurren

Inter arrival time: 913, 727, 015, 948, 309, 922, 753, 235, 302,

109, 93, 607, 738, 359, 888, 106, 21, 2, 493, 535.

Service time: 84, 10, 74, 53, 17, 79, 91, 67, 89, 38, 32, 94, 79, 05

79, 84, 52, 55, 30, 50.

The service time is distributed as follows:

Service time (min.)	Prob.	Cummulative prob.	random digit Assignment
1	0.10	0.10	01 - 10
2	0.20	0.30	11 - 30
3	0.30	0.60	31 - 60
4	0.25	0.85	61 - 85
5	0.10	0.95	86 - 95
6	0.05	1	96 - 00

Distribution of time b/w arrivals

(Interarrival) Time b/w arrivals	Prob	Cumulative prob.	Random Digit Assignment
1	0.125	0.125	001 – 125
2	0.125	0.250	126 – 250
3	0.125	0.375	251 – 375
4	0.125	0.500	376 – 500
5	0.125	0.625	501 – 625
6	0.125	0.750	626 – 750
7	0.125	0.875	751 – 875
8	0.125	1	876 – 000

Customer no.	Random no. for interarrival time	Interarrival in min	Arrival time	Random no. for service time	Service time	Time service begins	Time service ends	Time spent in system	Idle time	Wait
1	.		0	84	4	0	4	4	0	0
2	913	8	8	10	1	8	9	1	4	0
3	727	6	14	74	4	14	18	4	5	0
4	15	1	15	53	3	18	21	6	0	3
5	948	8	23	17	2	23	25	2	2	0
6	309	3	26	79	4	26	30	4	1	0
7	922	8	34	91	5	34	~~40~~ 39	5	4	0
8	753	7	41	67	4	41	45	4	1	0
9	235	2	43	89	5	45	50	7	0	2
10	302	3	46	38	3	50	53	7	0	4
11	109	1	47	32	3	53	56	9	0	6
12	93	1	48	94	5	56	61	13	0	8
13	607	5	53	79	4	61	65	12	0	8
14	738	6	59	5	1	65	66	7	0	6
15	359	3	62	79	4	66	70	8	0	4
16	888	8	70	84	4	70	74	4	0	0
17	106	1	71	52	3	74	77	6	0	3
18	212	~~1~~ 2	73	55	3	77	80	7 8	0	4 6
~~19~~	~~2~~	~~1~~	~~73~~	~~30~~	~~32~~	~~80~~	~~832~~	~~9 12~~	~~0~~	7
20	493	4	77	~~80~~ 30	3 2	8 ~~30~~	86 2 5	5	0	3 6

535	5	82	50	3	82	85	3	0	0	

Avg. waiting time ⇒ 51/20 = 2.55

" = (Total time customers wait / Total no. of customers)

Avg. Prob. of customers waiting = (no. of customers who wait) / Total no. of customers.

= 11/20 = 0.55.

Prob. of idle server = (total time idle) / total runtime of simulatn.

= 18/85 = 0.2117647

Avg. service time = total service time / no. of customers.

= 67/20 = 3.35.

Avg. time b/w arrival = Total time b/w arrivals / (no. of arrivals-1)

= 82/19 = 4.315789.

Avg. waiting time for those who wait = 51/11

= 4.636363

Avg. time customer spends in system = 118/20 = 5.9.

MS

Newspaper Seller's Problem.

→ Papers seller ~~cost~~ buys papers for 33 cents each and sells them for 50 cent each (loss profit from excess demand is 17 cents for each paper demanded that couldn't be provided)

→ newspapers not sold at the end of the day are sold as scrap for 5 cents each (salvage value of scrap papers)

→ they can be purchased in bundles of 10. Thus paper seller can buy 50, 60 and so on.

– There are 3 types of news days, 'good', 'fair' & 'poor' with probabilities of 0.35, 0.45 and 0.20 resp.

→ Problem is to determine optimal no. of papers the newspaper seller should purchase. – will be accomplished by simulating demands for 20 days

Demand Probability distribution

Demand	Good	Fair	Poor
40	0.03	0.10	0.44
50	0.05	0.18	0.22
60	0.15	0.40	0.16
70	0.20	0.20	0.12
80	0.35	0.08	0.06
90	0.15	0.04	0.00
100	0.07	0.00	0.00

M N Inventory system.

Suppose that max inventory level, M, is 11 units & the review period, N, is 5 days. Problem is to estimate by simulation, avg. ending units in inventory & no. of days when a shortage condition occurs.

Distributⁿ of no. of units demanded per day:

Demand	Probability	Cumulative prob.	random digit
0	0.1	0.1	01 – 10
1	0.25	0.35	11 – 35
2	0.35	0.7	36 – 70
3	0.21	0.91	71 – 91
4	0.09	1	92 – 00

Lead time → random variable.

Lead time (in days)	Probability.	cumulative prob	random digit.
1	0.6	0.6	01-6
2	0.3	0.9	7-9
3	0.1	1	0

Assume that orders are placed at close of business & are received for inventory at the beginning of business as determined by lead time. Simulate only 5 cycles. Simulation is started with inventory level at 3 units and order of 8 units scheduled to arrive in 2 days time.

Random digits (RD) for demand : 24, 35, 65, 81, 54, 03, 87, 27, 73, 70, 47, 45, 48, 17, 09, 42, 87, 26, 36, 40, 07, 63, 19, 88, 94

RD for lead time : 5, 0, 3, 4, 8

cycle	day	beginning inventory	RD for demand	Demand	Ending inventory	shortage quant.	order quant.	RD for lead time	days until order arrives.
1	1	3	24	1	2	0	-	-	1
	2	2	35	1	1	0	-	-	0
	3	1+8 = 9	65	2	7	0	-	-	-
	4	7	81	3	4	0	-	-	-
	5	4	54	2	2	0	9	5	1
2	1	2	03	0	2	0	-	-	0
	2	2+9 = 11	87	3	8	0	-	-	-
	3	8	27	1	7	0	-	-	-
	4	7	73	3	4	0	-	-	-
	5	4	70	2	2	0	9	0	3
3	1	2	47	2	0	0	-	-	2
	2	0	45	2	0	2	-	-	1
	3	0	48	2	0	4	-	-	0
	4	9	17	1	4	0	-	-	-
	5	4	09	0	4	0	7	3	1

4	1	4	42	2	2	0	-	-	0
	2	9	87	3	6	0	-	-	-
	3	6	26	1	5	0	-	-	-
	4	5	36	2	3	0	-	-	-
	5	3	40	2	1	0	10	4	1
5	1	1	07	0	1	0	-	-	0
	2	11	63	2	9	0	-	-	-
	3	9	19	1	8	0	-	-	-
	4	8	88	3	5	0	-	-	-
	5	5	94	4	1	0	10	8	2

reliability problem.

large milling machine has 3 different bearings that fail in service. The cumulative distributⁿ functⁿ of life of each bearing is identical.

bearing lifetime	1000	1100	1200	1300	1400	1500	1600	1700	1800	1900
PD	0.10	0.13	0.25	0.13	0.09	0.12	0.02	0.06	0.05	0.05
C.D.	0.10	0.23	0.48	0.61	0.7	0.82	0.84	0.~~88~~ 90	0.~~94~~ 95	~~0.99~~ 1
R.D.	1-10	11-23	24-48	49-61	62-70	71-82	83-84	85-90	91-95	96-00

when bearing fails, mill stops & repairperson is called and new bearing is installed. Delay time of repairperson arriving to millis a random variable with distribution.

Delay time	5	10	15
PD.	0.6	0.3	0.1

	C.D.	R.D.
5	0.6	1-6
10	0.9	7-9
15	1	0

downtime for mill is estimated at $5 per minute. Direct on-site cost of the repairperson is $ 15 per hour. It takes 20 min. to change one bearing, 30 min to change 2 bearings & 40 min to change three bearings. The bearings cost $16 each. simulate system for 20,000 hrs.

RD for lifetime for B1	67	08	49	84	44	30	10	63	2	2	77	59	23	53	85	75	45
RD for lifetime for B2	71	43	86	93	83	44	19	51	49	12	50	9	44	46	40	52	53
〃 B3	76	65	62	96	65	56	11	86	57	49	36	44	94	78	63	24	28
RD	3	7	5	1	4	3	7	8	8	3	2	4	1	6	2	7	0

Assumption: all bearings changed when a bearing fails.

B1 RD	B1 lifetime	B2 RD	B2 lifetime	B3 RD	B3 lifetime	smallest Accumulated lifetime	RD	RD delay
67	1400	71	1500	76	1500	1400	3	5
8	1000	43	1200	65	1400	+1000 2400	7	10
49	1300	86	1700	62	1400	+1300 3700	5	5
84	1600	93	1800	96	1900	+1600 5300	1	5
44	1200	83	1600	65	1400	+1200 6500	4	5
30	1200	44	1200	56	1800	+1200 7700	3	5
10	1000	19	1100	11	1100	+1000 8700	7	10
63	1400	51	1300	86	1700	+1300 10000	8	10
2	1000	49	1300	57	1300	+1000 11000	8	10
2	1000	12	1100	49	1300	+1000 12000	3	5
77	1500	50	1300	36	1200	+1200 13200	2	

20,900

Other simulation problem.

standard normal variate $Z = \frac{X - \mu}{\sigma}$ $X = Z\sigma + \mu$

Daily Demand	Prob.	C.P.	R.D
3	0.20	0.20	~~01~~ 01 - 20
4	0.35	0.55	21 - 55
5	0.30	0.85	56 - 85
6	0.15	1	86 - 00

Lead time	Prob.	C.P.	R.D
1	0.36	0.36	01 - 36
2	0.42	0.78	37 - 78
3	0.22	1	79 - 00

To find lead-time demand, make the table, assign R.D. (lead time) to value & acc. to value of lead time, assign that many R.D (demand) and find values & summation of it.

Random var. X, Y, Z are distributed as follows:

$X \sim N(\text{mean} = 100, \text{variance} = 100)$

$Y \sim$ —(" — $= 300$, —" — $= 225)$

$Z \sim$ —(" — $= 40$, —" — $= 64)$

Simulate 5 values of random variable

$W = (X+Y)/Z$

RNNx = −0.137, −0.918, 1.692, −0.199, −0.411

RNNy = 0.577, 0.303, ~~0~~, −0.383, 1.033, 0.633

RNNz = −0.568, −0.384, −0.198, 0.031, 0.397

X	Y	Z
98.63	308.655	
90.82	304.545	
116.92	294.255	
98.01	315.495	
95.89	309.495	

$Z = \frac{X - \mu}{\sigma}$

For X ⇒ $Z = \frac{X - 100}{10}$

$X = 10Z + 100$

For Y ⇒ $Z = \frac{Y - 300}{15}$

$Y = 15Z + 300$

For Z ⇒ $Z^{*} = \frac{Z^{*} - 40}{8}$

$Z^{*} = 8Z + 40$

MS

→ Uniformity: Freq. test.
↳ KS test ↳ Chi square test

→ Independence: Runs test, Autocorrelation test, gaps test, Pokers Test.

Testing for independence:

$H0$: $R_i \sim$ independently

$H1$: $R_i \nsim$ independently.

Testing for uniformity:

$H_0: R_i \sim U[0,1]$

$H_1: R_i \nsim U[0,1]$

Type I error(α) $\alpha = P(\text{reject } H_0 \mid H_0 \text{ is true})$

α is set 0.01 or 0.05

Type II error (β)

Types of tests: Theoretical tests & Emperical tests.

type 1 error → reject H_0 when H_0 is true

type 2 error → failing to reject H_0 when ~~alternative~~ H_1 is true

Kolmogorov- Smirnov Test.

$S_N(x)$ → N sample observations

~~CDF~~ Continuous CDF $F(x) = x$, $0 \leq x \leq 1$

RN generated samples ⇒ $R_1, R_2 \text{---} R_N$

Empirical CDF, $S_N(x)$

$$S_N(x) = \frac{\text{no. of } R_1, R_2 \text{---} R_N \text{ which are} \leq x}{N}$$

Based on statistic ~~stat~~ $D = \max |F(x) - S_N(x)|$

steps:

1: Define hypothesis for testing uniformity
2: Arrange data in ascending order
3: compute $D^+ = \max \{ i/N - R_{(i)} \}$

$D^- = \max \{ R_{(i)} - (i-1)/N \}$

4: compute $D = \max(D^+, D^-)$
5: Determine the critical, D_α for specified significance level α and given sample size N from Table.
6: If $D > D_\alpha \Rightarrow H_0$ is rejected ; null hypothesis that the data are a sample from uniform distributⁿ is rejected.
Else ~~then~~ no diff. has been detected b/w sample distributⁿ & uniform distributⁿ.

eg: nos. ⇒ 0.44, 0.81, 0.14, 0.05, 0.93.

$\alpha = 0.05$

use KS test to determine if hypothesis that the nos. are uniformly distributed on the interval [0, 1] can be rejected.

step 1: Defining hypothesis

$H_0 : R_i \sim U[0,1]$

$H_1 : R_i \not\sim U[0,1]$

step 2: 0.05, 0.14, 0.44, 0.81, 0.93

step 3:

R_i	0.05	0.14	0.44	0.81	0.93
i/N	1/5 = 0.20	2/5 = 0.40	3/5 = 0.60	4/5 = 0.80	5/5 = 1
$i/N - R_i$	0.15	0.26	0.16	–	0.07
$(R_i - (i-1)/N)$	0.05	–	0.04	0.21	0.13

$D^+ = \max\{i/N - R_i\}$

$= 0.26$

$D^- = \max\{ R_i - (i-1)/N \}$

$= 0.21$

step 4: $D = \max(D^+, D^-)$

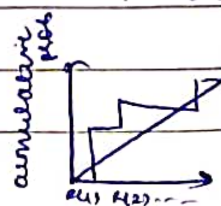
$= 0.26$

step 5: For $\alpha = 0.05$, $N = 5$

$D_\alpha = 0.565$

step 6: $D_\alpha > D$

∴ H_0 is not rejected



comparison of F(x) & $S_N(x)$

0.63, 0.49, 0.24, 0.89, 0.57, 0.71

$\alpha = 0.05$

$D_\alpha = 0.521$ for $N = 6$.

Chi Square Test.

Define the hypothesis: $H_0 : R_i \sim U[0,1]$

$H_1 : R_i \nsim U[0,1]$

consider seq. of R.N., we divide it into n classes of equal length

Let O_i be values that fall within an interval (i^{th} class).
↳ observed values.

$E_i \rightarrow$ expected no. of no. in i^{th} class.

For uniform distributⁿ, $E_i \rightarrow$ expected no. in each class.

$E_i = N/n$; $N \Rightarrow$ total no. of observations.

Compute sample test statistics

Determine critical value for specified significance level and degrees of freedom

If $X_0^2 > X_{\alpha n-1}^2 \Rightarrow$ H_0 is rejected, else no difference has been detected b/w the sample distribution and uniform distributⁿ.

$$X_0^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$ for large sample $N \geq 50$

n - 1 degrees.

~~Data must be random sample~~

$H_0 : R_i \sim U[0,1]$

$H_1 : R_i \nsim U[0,1]$

$\alpha = 0.05$

$n = 10$

$[0, 0.1), [0.1, 0.2) \text{ ---- } [0.9, 1.0)$

P.T.O

class distributⁿ	O_i	E_i (N/n)	$O_i - E_i$	$(O_i-E_i)^2$	$(O_i-E_i)^2/E_i$
0 - 0.1	7	10	-3	9	0.9
0.1 - 0.2	9	10	-1	1	0.1
0.2 - 0.3	8	10	-2	4	0.4
0.3 - 0.4	9	10	-1	1	0.1
0.4 - 0.5	14	10	4	16	1.6
0.5 - 0.6	7	10	-3	9	0.9
~~0.6 - 0.7~~	10	10	0	0	0
0.7 - 0.8	15	10	5	25	2.5
0.8 - 0.9	9	10	-1	1	0.1
0.9 - 1.0	12	10	2	4	0.4
					$\Sigma \Rightarrow 7$

$\chi_0^2 = 7$ $\qquad \chi_0^2 < \chi^2_{0.05,9}$

$\chi^2_{0.05,9} = 16.9$

∴ null hypothesis is not rejected.

→ Ho is accepted. (not rejected)

Independence

→ Runs test
- up & down
- above & below the mean
- length of runs
  - up & down
  - above & below mean

Examines arrangement of nos. in seq. to test hypothesis of independence.

Ho: R_i ~ independently

H1: R_i ≁ not independently.

Seq. of tossing coin: (H)(T T)(H H)(T T T)(H)(T) → 6 runs

3 mutually exclusive events → head, tail, no event

Length: 1^{st} : 1

2^{nd} : 2

3^{rd} : 2

4^{th} : 3

5^{th} : 1

6^{th} : 1

0.87, 0.15, 0.23, 0.45, 0.69, 0.32, 0.30, 0.19, 0.24, 0.18, 0.65, 0.82, 0.93, 0.22, 0.81

upward run: no. followed by larger no.

down run: —"— smaller no.

− + + + − − − + − + + + − +

Runs = 8

length = 1, 3, 3, 1, 1, 3, 1, 1

How many runs? what is the length.

→ 0.08, 0.18, 0.28, 0.31, 0.45, 0.54, 0.63, 0.73, 0.86, 0.91

Runs = 1 ; length = 9

For N > 20, a approx. by normal distributⁿ, $N(\mu_a, \sigma_a^2)$

$$Z_0 = \frac{(a - \mu_a)}{\sigma_a}$$; a ⇒ total no. of runs.

$$\mu = \frac{2N-1}{3} \qquad \sigma^2 = \frac{16N-29}{90}$$

Acceptance ~~range~~ region for hypothesis of independence:

$$-Z_{\alpha/2} \leq Z_0 \leq Z_{\alpha/2}$$

Define hypothesis for testing the independence:

H_0: R_i are independently

H_1: $R_i \neq$ independently.

Write down seq. of runs up & down.

Compute mean & variance of 'a'

$$\mu = \frac{2N-1}{3}, \quad \sigma_a^2 = \frac{16N-29}{90}$$

compute standard normal statistics.

$Z_0 = (a - \mu_a) / \sigma_a$

Determine critical value $-Z_{\alpha/2}$ and $Z_{\alpha/2}$ for specified significance level α ~~for~~ from table A.3

If $-Z_{\alpha/2} \le Z_0 \le Z_{\alpha/2}$

→ H_0 is not rejected for significance level α.

Based on runs down, det. ~~that~~ whether foll seq. of 40 nos. is such that hypothesis of independence can be rejected where $\alpha = 0.05$ (critical value is 1.96.)

0.41	0.68	0.89	0.94	0.74	0.91	0.55	0.62	0.36	0.27
0.19	0.72	0.75	0.08	0.54	0.02	0.01	0.36	0.16	0.28
0.18	0.01	0.95	0.69	0.18	0.47	0.23	0.32	0.82	0.53
0.31	0.42	0.73	0.04	0.83	0.45	0.13	0.57	0.63	0.29

+ + + − + − + − − − + + − + − − + − + − − + − − + − + + − −

+ + − + − − + + −

Runs = 26

$a = 26$

H_0: $R_i \sim$ independently.

H_1: $R_i \not\sim$ independently.

$N = 40$

$$\mu = \frac{2N-1}{3} = \cancel{26.33}\ \frac{79}{3} = 26.33$$

$$\sigma_a^2 = \frac{16N - 29}{90} = \frac{611}{90}$$

$$\sigma_a = 2.605$$

$$Z_0 = \frac{a - \mu_a}{\sigma_a}$$

$$= \frac{26 - \cancel{26.33}\ 26.333}{3(2.605)} = \frac{\cancel{1}3}{\cancel{7.815}} = -0.127$$

$-Z_{\alpha/2} \le Z_0 \le Z_{\alpha/2}$

$-1.96 \le Z_0 \le 1.96.$

$Z_0 = -0.127$

$\therefore$ It is within range

Ho is accepted. (Ho is not rejected for significance level 0.05)

runs above & below the mean → + when value is above the mean, - when value is below the mean.

$$\mu_b = \cancel{(2n_1n_2}\ \frac{(2n_1n_2)}{N} + \frac{1}{2}$$

and $$\sigma_b^2 = \frac{2n_1n_2(2n_1n_2 - N)}{N^2(N-1)}$$

If Z_0 is b/w $-Z_{\alpha/2} \le Z_0 \le Z_{\alpha/2}$, hypothesis accepted

$$Z_0 = \frac{b-\mu}{\sigma}$$

n_1 → no. of observations above the mean

n_2 → — " — below — " —

$Z_{critical} = 1.96$

mean ⇒ .00 ⇒ $\frac{0.99}{2} = 0.495$

.99

$b = 17$

− + + + + + + + + − − − $n_1 = 18$

− + + − + − − − − − $n_2 = 22$

− − + + − − − − + + $N = 40$

− − + − + − − + + −

$$\mu_b = \frac{18 \times 22 \times 2}{40} + \frac{1}{2} = \frac{99}{5} + \frac{1}{2} = 19.8 + 0.5 = 20.3$$

$$\sigma_b^2 = \frac{(18 \times 22 \times 2)((18 \times 22 \times 2) - 40)}{40^2(39)}$$

$$\frac{(792)(752)}{1600(39)} = 9.5446$$

$\sigma_b = 3.0894$

$$Z_0 = \frac{b-\mu}{\sigma}$$

$$= \frac{17 - 20.3}{3.08984}$$

$= -1.068$

$\therefore$ Ho is accepted

$\frac{3102}{325}$

classmate

Date ______
Page ______

- Length of runs → up & down
  → above & below the mean

Define hypothesis for testing for independence as

Ho: Ri ~ independently

H1: Ri ≁ "

Write down the sequence of runs up and runs down or runs above & runs below the mean (as per problem statement)

Find length of the runs in the sequence.

Prepare table of runs in the sequence for no. of observed runs of each length

Compute the expected value of Y_i where Y_i is no. of runs length i in the seq. of n nos.

$$E(Y_i) = \frac{2}{(i+3)!}\left[N(i^2+3i+1) - (i^2+3i-i-4)\right] \quad ; i \leq N-2$$

$$E(Y_i) = \frac{2}{N!} \quad ; \cancel{i < N-1} \quad i = N-1$$

above & below the mean: $E(Y_i) = \frac{N w_i}{E(I)}$, $N > 20$

where w_i = approx. prob. that a run has length i and is given by $w_i = \left(\frac{n_1}{N}\right)^i \left(\frac{n_2}{N}\right) + \left(\frac{n_1}{N}\right)\left(\frac{n_2}{N}\right)^i$, $N > 20$

$E(I)$ is approx. expected length of a run and is given by

$$E(I) = \frac{n_1}{n_2} + \frac{n_2}{n_1}, \quad N > 20$$

Compute the mean or expected total no. of runs in a sequence

For runs up & down: $\mu_a = \frac{2N-1}{3}$

For runs above & below the mean: $E(A) = \frac{N}{E(I)}$, $N \geq 20$

Compute expected no. of ~~length~~ runs of length greater than or equal to max. of length of observed run.

ii) For runs up & down : $\mu_a = \sum_{i=1}^{m} E(Y_i)$

where $m \to$ max. length of observed runs.

~~~~ iii) For runs ~~~~ above & below the mean :

$$E(A) = \sum_{i=1}^{m} E(Y_i)$$

where $m \to$ max. length of observed runs.

Apply chi square test.

The test statistics is given by : $\chi_0^2 = \sum_{i=1}^{n} \frac{[O_i - E(Y_i)]^2}{E_i}$

$n \Rightarrow$ no. of classes

Determine critical value for specified significance level α with $n-1$ degrees of freedom

If ~~chi square~~ $\chi_0^2 < \chi^2_{\alpha, n-1} \to H_0$ is not rejected (accepted)

$\alpha = 0.05$

Runlength	Obs. no. of runs O_i	Expected no. of runs $E(Y_i)$	$\frac{(O_i - E(Y_i))^2}{E_i}$

Define hypothesis :

H_0 : $R_i \sim$ independently

H_1 : $R_i \nsim$ independently

0.30 0.48

0.48

0.42

0.95

0.73

0.60
~~~~

classmate

Date ____

Page ____

+ − − + − + + + + +   Runs = 40

− + − + + − + + − +

− + − − + + + − − +

− − + + + − − − + +

− − − + + + − − − − +

− + − − − + − + + −

Runlength (i)	~~Obs~~ O_i	$E(Y_i)$		$(O_i - E(Y_i))^2 / E_i$
1	26	25.083		0.03
2	9 } 14	10.~~8123~~ 766	} 14.59	0.02
3	5	3.82		
	40	39.67		0.05

1) $E(Y_i) = \frac{2}{(i+3)!}\left[N(i^2+3i+1) - (i^3+3i^2-i-4)\right]$

$= \frac{2}{4!}\left[60(1+3+1) - (1+3-1-4)\right]$

$= \frac{602}{24} = 25.083$

2) $E(Y_i) = \frac{2}{5!}\left(60(4+6+1) - (8+12-2-4)\right]$

$= \frac{2(660-14)}{120} = \frac{646}{60} = 10.766$

3) $E(Y_i) = \frac{2}{6!}\left(60(9+9+1) - (27+27-3-4)\right)$

$= \frac{2(1140-47)}{720} = \frac{1093}{360} = 3.036$

critical value $\chi^2_{0.05,1} = 3.84$