~~The test is no.~~

Gap Test

Measures no. of digits b/w successive occurences of the same digit

Every probability $(0.9)^{10}(0.1)^n$

* Specify the CDF for the theoretical frequency distribution given by

$$F(\text{gaps} <= x) = F(x) = 0.1 \sum_{n=0}^{x} (0.9)^n =$$

Digits length

2	3	6	5	6	0	0	1	3	4
5	6	7	9	4	9	3	1	8	3
1	3	7	4	8	6	2	5	1	6
4	4	3	3	4	2	1	5	8	7
0	8	8	2	6	7	8	1	3	5
3	8	4	0	9	0	3	0	9	2
4	6	9	9	8	5	6	0	1	7
6	7	0	3	1	0	2	4	2	0
1	1	2	6	7	6	3	7	5	9
3	6	6	7	8	2	3	5	9	6
6	4	0	3	9	3	6	8	1	6
0	7	6	2	6	0	5	7	8	0

Digits	No. of Gaps	length of Gaps.
0	13	0, 33, 12, 1, 1, 9, 4, 2, 3, 22, 7, 4, 3
1	10	9, 10, 27, 7, 10, 20, 5, 5, 0, 26
2	9	25, 18, 7, 15, 16, 1, 3, 12, 17
3	15	6, 7, 2, 1, 10, 0, 14, 1, 5, 16, 12, 3, 5, 6, 1
4	9	4, 8, 6, 0, 2, 17, 7, 16, 23,
5	9	6, 16, 9, 11, 15, 22, 8, 11, 6
6	17	1, 6, 13, 3, 14, 16, 4, 3, 12, 1, 5, 0, 6, 0, 5, 5, 1
7	10	9, 15, 5, 23, 1, 12, 2, 5, 17, 5
8	10	5, 13, 2, 0, 3, 4, 12, 29, 12, 10
9	8	1, 38, 3, 3, 0, 25, 8, 5

Total number of gaps = no of digits − Number of distinct digits

= 120 − 10 = 110

Gap length	Freq	Relative Freq	F(x) CF	S(x) Therotical CF $1-(0.9)^{x+1}$	$\lvert F(x)-S(x) \rvert$
0-3	34	0.3091	0.3091	0.3439	0.0348
4-7	30	0.2727	0.5818	0.5695	0.0123
8-11	13	0.1182	0.7	0.7176	0.0176
12-15	13	0.1182	0.8182	0.8147	0.0035
16-19	9	0.0818	0.9	0.8784	0.0216
20-23	5	0.0455	0.9455	0.9202	0.0253
24-27	3	0.0273	0.9728	0.9477	0.0251
28-31	1	0.0091	0.9819	0.9657	0.0162
32-35	1	0.0091	0.991	0.9775	0.0135
36-38	1	0.0090	1	0.9836	0.0164

$t_{calc} < t_\alpha$

$\frac{1.36}{\sqrt{10}} = 0.43$

H0 is accepted.

~~General RTR~~

what so called	Delay	Activity
	a conditional wait	an unconditional wait
A completion	a secondary event	a primary event
A management	represented in FEL	represented in the list one after other

Clock t	System state $(x, y, z \ldots)$	Entities & Attributes	Set 1	Set 2	..	Future event list FEL	Cumulative Statistics & Counters
						$(3, t_1)$ - Type 3 event occurring at time t_1	
						$(1, t_2)$ - Type 2 event occurring at time t_2	
						⋮	

world views:
- → Event-scheduling
- → process-interaction
- → activity-scanning

Event-Scheduling / Time-advanced Algorithm.

→ The mechanism for advancing simulation time & guaranteeing that all events occur in correct chronological order is based on FEL

→ FEL

$$t < t_1 \leq t_2 \leq \ldots \leq t_n$$

current value of simulation time

imminent event

→ Scheduling a future event list.

- LQ(t) - no. of customers in waiting line
- LS(t) - no. of customers being served at time t
- Event notices (event type, event time)
- Activities:
  Interarrival Time
  Service Time
- Delay : Customer time spent in waiting line.

Q) t= 60

Interarrival Time	8, 6, 1, 8, 3.
Service Time	4, 1, 4, 3, 2, 4

Event Scheduling Algorithm.

Clk	System State LQ(t)	LS(t)	FEL [t+sn] [t+an]	Busy Statistics (cumulative
0	0	1	(D, 4); (A,8); (E,60)	0
4	0	0	(A,8); (E,60)	4
8.	0	1	(E,60) (D, 9); (A,14)	4
9	0	0	(A,14) (E,60)	5
14	0	1	(A,15) (D,18) (E,60)	5.
15	1	1	(A,23) (D,18) (E,60)	6
18	0	1	(A,23) (E,60) (D,21)	9
21	0	0	(A, 23) (E, 60)	12.

Q) Inter Arrival Time : 1, 1, 6, 3, 7, 5, 2, 4, 1, 5

Service Time : 5, 3, 5, 5, 2, 6, 5, 2, 4, 4, 6

Determine: (i) The sum of customer response time for all the customers who have departed until the current simulation time

(ii) The total no. of customers who spend 4 or minutes in the system

(iii) Total no. of customers who have departed until the current simulation time.

Event Scheduling Algorithm

CLK	System state LQ(t)	LS(t)	FEL [t+s^n] [t+a^n]	Response Time	Waiting Time
0	0	1	(D,5) ; (A,1) ; (E,10)	0	0
1	1	1	(D,5) ; (A,2) ; (E,10)		
2	2	1	(D,5) ; (A,8) ; (E,10)		
5	1	1	(D,8) ; (A,8) ; (E,10)		
8	0	1	(D,13) ; (A,8) ; (E,10)		
8	1	1	(D,13) ; (A,11) ; (E,10)		

Simulation

Process of designing a model of a real system & conducting experiments with the model for the purpose of understanding the behavior for the system.

- A duplication of the original system
- To understand the implimentation of the system.

Monte-Carlo Simulation

- It is an Experiment on chance
- Uses random number & requires decision making under uncertainty

Understanding of this technique

To understand their technique, this is break down into 5 steps:

1. Establishing Probability Distribution
2. Cumulative Probability Distribution
3. Setting Random Number Intervals
4. Generating Random Numbers.
5. To find the answer of question asked using the above 4 steps.

Q)

Category	Time Required	No. of Patients	Probability	CP	RNO
Filling	45 min	40	0.4	0.4	00-39
Crown	60 min	15	0.15	0.55	40-54
Cleaning	15 min	15	0.15	0.70	55-69
Extracting	45 min	10	0.1	0.8	70-79
Checkup	15 min	20	0.2	1.0	80-99

Simulate the dentist's clinic four hours & find ~~the~~ out the average waiting time starting at 8.00 am as well as idleness of doctor. 30 mins slot

Random digits: 40, 82, 11, 34, 25, 66, 17, 79

Patient No.	AT	Random Digit	Category	ST
1	8.00	40	Crown	60
2	8.30	82	Checkup	15
3	9.00	11	Filling	45
4	9.30	34	Filling	45
5	10.00	25	Filling	45
6	10.30	66	Cleaning	15
7	11.00	17	Filling	45
8	11.30	79	Extracting	45

Patient No.	AT	Service Start	Service Duration	Service End	Waiting min	Idle Time
1	8.00	8.00	60	9.00	0	0
2	8.30	9.00	15	9.15	30	0
3	9.00	9.15	45	10.00	15	0
4	9.30	10.00	45	10.45	30	0
5	10.00	10.45	45	11.30	45	0
6	10.30	11.30	15	11.45	60	0
7	11.00	11.45	45	12.30	45	0
8	11.30	12.30	45	1.15	60	0
					$\Sigma = 285$	

Avg. waiting time $= \frac{285}{8} = 35.625$ min

Avg. Idle time $=$ 0 min

Dump-Truck Problem:

Loading Time	10	5	5	10	15	10	10
Weighing Time	12	12	12	16	12	16	
Travel Time	60	100	40	40	80		

Clk	System State				Lists			Statistics	
	LQ(t)	L(t)	WQ(t)	W(t)	Loader Q	Weigh Q	FEL	BL	BS
0	3	2	0	1	DT4, DT5, DT6	0	(EW, 12, DT1) (EL, 10, DT2) (EL, 5, DT3)	0	0
5	2	2	1	1	DT5, DT6	DT3	(EW, 12, DT1) (EL, 10, DT2) (EL, 10, DT4)	2×5	1×
10	1	2	2	1	DT6	DT3 DT2	(EW, 12, DT1) (EL, 10, DT4) (EL, 20, DT5)	20	
10									

Probability & Statistics in Simulation :

Study of random phenomenon, by constructing ideal probabilistic model of real world simulation.

Ideal - consists of all possible outcomes & corresponding probabilities

Role is to analyse the behaviour of system as per the given distribution & contributed probabilities.

Need :

- To validate the simulation model
- To determine / choose the input probability distributions
- To perform statistically analyze the output
- Needed to design / construct.

Q. variance of {75, 90, 40, 95, 80}

$$V(X) = E(X^2) - [E(X)]^2$$

$$= \frac{30750}{5} - (76)^2$$

$\frac{\Sigma p_i x_i}{\Sigma x_i}$

$$= 6150 - 5776$$

$$= 374$$

Q. Consider x to be a random variable having mass function

$$P(x) = x/6, \quad x = 1, 2, 3$$
$$= 0, \quad \text{otherwise}$$

compute (i) mean (ii) variance (iii) standard deviation.

~~Mean = 2/6~~

~~Var~~ Mean $= \frac{1\left(\frac{1}{6}\right) + 2\left(\frac{2}{6}\right) + 3\left(\frac{3}{6}\right)}{\cancel{6}} = \frac{1+4+9}{36} = \frac{14}{36}$

$$\text{Variance} = \frac{\left(\frac{14}{18}-\frac{1}{6}\right)^2 + \left(\frac{14}{18}-\frac{2}{6}\right)^2 + \left(\frac{14}{18}-\frac{3}{6}\right)^2}{6}$$

$$= \cancel{\frac{\left(\frac{4}{18}\right)^2 + \left(\frac{1}{18}\right)^2 + \left(\frac{-2}{18}\right)^2}{6}}$$

$$= \cancel{\frac{16+1+4}{1944}}$$

$$= \cancel{\frac{22}{1944} = \frac{11}{972}} = \cancel{0.0}$$

$$= \left(\frac{13}{6}\right)^2 + \left(\frac{14}{6}\right)^2 + \left(\frac{11}{6}\right)^2$$

$$= \frac{434}{36}$$

$$= \frac{217}{18} = 12.0556$$

$$SD = 3.4\cancel{42}$$

$$\text{variance} = E(x^2) - [E(x)]^2$$

$$= 1\left(\tfrac{1}{6}\right)+4 \qquad \left(\frac{14}{6}\right)^2$$

$$=$$

$$= 5.445$$

Page

Bernoulli distribution

$E(X) = p$

$V(X) = pq$

Binomial Distribution

$$p(x) = \begin{cases} nC_x \; p^x q^{n-x} & x = 0,1,2,\ldots n \\ 0 & \text{otherwise} \end{cases}$$

$$nC_x = \frac{n!}{x!\,(n-x)!}$$

$E(X) = np$

$V(X) = npq$

Eg: The trait of having a klinkled flooje occurs on 10% of kreptoplompans.
20 " , what is probability of 3 of them?

$P(3) = {}^{20}C_3 \, (0.1)^3 \, (0.9)^{17}$

$= 0.1902$

Eg: The prob. of computer chip failure 0.05. Everday a random sample of 14 is taken. what is the probability that
i) atmost 3 chips will fail
ii) atleast 3 chips will fail

$P(X \leq 3) = 0.996$

$P(X \geq 3) = 0.303$

Geometric Distribution

$$p(x) = \begin{cases} q^{x-1} p & x = 1, 2 \\ 0 & \text{otherwise} \end{cases}$$

$F(x) = 1 - q^x$

$E(x) = 1/p$

$V(x) = q/p^2$

Eg: what is the proba that the first kreptoplomian found to have a klinkled Hoge will be the 5th kreptopolomian overall?

$(0.9)^4 (0.1) = 0.0656$

Poisson Distribution

→ Number of events that occur in a given period (catch words)

$$P(X) = \begin{cases} \frac{e^{-a} a^x}{x!} & , X = 0, 1 \ldots \\ 0 & , \text{otherwise} \end{cases}$$

α is mean arrival time

$E(x) = V(x) = \alpha$

$$F(x) = \sum_{i=0}^{x} \frac{e^{-a} a^i}{i!}$$

Date
Page

Continuous Random variables

Eg: The life of a laser ray device used to inspect cracks in an aircraft is given by x a continuous r.v. $x \geq 0$.
The pdf is given by : $f(x) = \frac{1}{2} e^{-x/2}$ where $x \geq 0$

$= 0$ otherwise

Find the probability that lifetime is b/w 2-3 years.

$$\int_2^3 \frac{1}{2} e^{-x/2}$$

$$\frac{1}{2(-1/2)} \int_2^3 e^{-x/2}$$

$$-e^{-x/2}$$

$$\cancel{(e^{-3/2}} = -e^{-\left(\frac{3-2}{2}\right)}$$

$$= -e^{-1/2}$$

$$= \cancel{-0}.\ 0.145$$

$$F(x) = \int_{-\infty}^{x} f(t)\, dt$$

$$= \int_{-\infty}^{0} f(t)\, dt + \int_0^x \frac{1}{2} \times e^{-x/2}$$

$$= \frac{1}{2 \times \left(-\frac{1}{2}\right)} \times \int_0^x e^{-x/2}$$

$$= -\{e^{-x/2} + 1\}$$

uniform distribution

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{if } a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

$$F(X) = \frac{x-a}{b-a}$$

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx = \frac{b+a}{2}$$

$$V(X) = \frac{(b-a)^2}{12}$$

Exponential Distribution:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

$$E(X) = \frac{1}{\lambda}$$

$$V(X) = \frac{1}{\lambda^2}$$

$$F(x) = \begin{cases} 0, & x < 0 \\ \int_0^x \lambda e^{-\lambda t}\,dt = 1 - e^{-\lambda x}, & x \ge 0 \end{cases}$$

Q) A passenger arrives at a railway stn. randomly which is uniformly distributed b/w 10am and 10:30 am. Find the probability that the passenger has to wait for the train for (i) < 6 mins ii) > 10 mins

Assume that the local trains arrive at the railway station every 15 mins beginning at 5 am.

Case 1: 9-15, 24-30.

case 2: 1-5, 15-20

Q) The component of a system whose time to failure is exponentially distribut with $\lambda = 1/6$. If 6 such components are installed in different systems, what is the probability that atleast 2 are still working at the end of 9 years

Normal distribution

$\mu = 0$ $\quad \sigma^2 = 1.$ $\quad N(0, 1)$

standard normal distribution

$X \sim N(\mu, \sigma^2)$

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

$$F(x) = \Phi\left(\frac{x - \mu}{\sigma}\right)$$

$\Phi(z)$ - cumulative density function

A.3 from text book.

Q) Student GPAs are approximately normally distributed with $\mu = 2.4$ & $\sigma = 0.8$. What fraction of students will possess a GPA in excess of 3.0?

$$Z = \frac{X - 2.4}{0.8}$$

where $X > 3.0$

hence $1 - F(3.0)$

$F(3.0) = \Phi[(3.0 - 2.4)/0.8] = \Phi[0.75]$

From table, 0.77337

Ans $= 1 - 0.77337 = 0.2266$

A class of 50 students have exam scores that are normally distributed with $\mu = 70$ & $\sigma = 9.8$

The teacher grades on a 'curve' with the following policy for score X

$X < \mu - 2\sigma \rightarrow F$

$\mu - 2\sigma < X < \mu - \sigma \rightarrow D$

$\mu - \sigma < X < \mu + \sigma \rightarrow C$

$\mu + \sigma < X < \mu + 2\sigma \rightarrow B$

$\mu + 2\sigma < X \rightarrow A$

No score will be exactly $\mu \pm k\sigma$ for any k.
How many students will get B grades

$F(\mu + 2\sigma) - F(\mu + \sigma)$

$F(\mu + 2\sigma)$ we get $Z = \frac{X - \mu}{\sigma} = \frac{\mu + 2\sigma - \mu}{\sigma} = 2$

$F(\mu + \sigma)$ we get $Z = \frac{(\mu + \sigma - \mu)}{\sigma} = 1.$

$0.97725 - 0.84134 = 0.13591$

Number of Bs $= 50(0.13591) = 6.7955$

Triangular Distribution

PDF

$$f(x) = 2(x-a) / (b-a)(c-a) \quad \text{if } a \le x \le b$$
$$= 2(c-x) / (c-b)(c-a) \quad \text{if } b \le x \le c$$
$$= 0 \quad \text{otherwise}$$

CDF

$$F(X) = 0 \quad \text{if } x \le a$$
$$= (x-a)^2 / (b-a)(c-a) \quad \text{if } a \le x \le b$$
$$= 1 - [(c-x)^2 / (c-b)(c-a) \quad \text{if } b \le x \le c$$
$$= 1 \quad \text{if } x > c$$

$$E(X) = \frac{a+b+c}{3}$$

$$var = \frac{a^2 + b^2 + c^2 - ac - bc - ab}{18}$$

Poisson Process

$$p(x) = \begin{cases} \frac{e^{-a} a^x}{x!}, & x = 0, 1, \ldots \\ 0, & \text{otherwise} \end{cases}$$

$$F(x) = \sum_{i=0}^{x} \frac{e^{-a} a^i}{i!}$$

Exponential Distribution for random number generation

$F(X) = 1 - e^{-\lambda x}$

$R \sim U[0, 1]$ $\quad R_1, R_2, R_3 \ldots$

$F(X) = R$

$1 - e^{-\lambda x} = R$

$$x_i = -\frac{1}{\lambda} \ln(1 - R_i)$$

For uniform distribution

$$F(x) = \frac{x-a}{b-a}$$

$R \sim U[0,1]$ $\quad R_1, R_2, R_3 \ldots$

$$F(x) = R$$

$$\frac{x-a}{b-a} = R$$

$$\Rightarrow x_i = R_i(b-a) + a$$

Q) A system has a component whose time to failure is exponentially distributed with failure rate $\frac{1}{6}$. Generate 2 random failure time from this distribution. Use $R_1 = 0.38$ and $R_2 = 0.45$

$$x_i = -\frac{1}{\lambda} \ln(1-R_i)$$

$$x_1 = -6 \ln(1-0.38)$$
$$= 2.868$$
$$x_2 = -6 \ln(1-0.45)$$
$$= 3.587$$

Q) The time required for a person to travel from place a to place b is uniformly distributed over the interval [30,35]. Generate 2 random travel times from this distribution. Use $R_1 = 0.5025$ $\quad R_2 = 0.2916$

$$x_1 = 0.5025 \times 5 + 30 = 32.5125$$
$$x_2 = 0.2916 \times 5 + 30 = 31.458$$

Q) A digital sensor is used to determine the quality of computer chips. The QC department rejects the chips those that fail to follow the triangular distribution with $a=0$ & $b=1$ & $c=3$

$$x_i = \sqrt{c R_i} \qquad 0 < R_i \le 1/c$$

$$x_i = c - \sqrt{c(c-1)(c-R_i)} \qquad 1/c < R_i \le 1$$

<u>Empirical continuous Distribution</u>

$$X = F^{-1}(R) = x_{(i-1)} + a_i\left(R - \frac{(i-1)}{n}\right)$$

$$a_i = \frac{x_{(i)} - x_{(i-1)}}{i/n - (i-1)/n} = \frac{x_{(i)} - x_{(i-1)}}{1/n}$$

When data set is <u>very small</u>.

<u>Grouped Data</u>

$$a_i = \frac{x_i - x_{(i-1)}}{c_i - c_{(i-1)}}$$

$$X = x_{(i-1)} + a_i(R - c_{(i-1)})$$

Q) Suppose the data collected for 100 broken widget repair times are.

i	Interval (Hours)	Frequency	RF (F/100)	CF	Slope (a_i)
1	$0.25 \le x \le 0.5$	31	0.31	0.31	0.81
2	$0.5 \le x \le 1.0$	10	0.10	0.41	5
3	$1.0 \le x \le 1.5$	25	0.25	0.66	2
4	$1.5 \le x \le 2.0$	34	0.34	1.00	1.47

$R_1 = 0.83$

$C_3 = 0.66 < R_1 < C_4 = 1.00$

$X_1 = X_{(4-1)} + a_4 (R_1 - C_{(4-1)})$

$= 1.5 + 1.47 (0.83 - 0.66)$

$= 1.75$

Q) 5 observations of response times to incoming alarms have been collected and are used in a simulation to investigate the scheduling policies for alternate staffing. The response times are: 1.60, 0.55, 1.12, 2.20, 1.94. Setup a table for generating response time by the table lookup method and generate 2 values of response times using uniform random numbers: 0.5426 and 0.1524

Ascending order → 0.55, 1.12, 1.60, 1.94, 2.20

Interval	Probabily	CF
$0 < X \le 0.55$	0.2	0.2
$0.55 < X \le 1.12$	0.2	0.4
$1.12 < X \le 1.60$	0.2	0.6
$1.60 < X \le 1.94$	0.2	0.8
$1.94 < X \le 2.20$	0.2	1.0

$R_1 = 0.5426$
$X_1 = 1.46224$

$R_2 = 0.1524$
$X_2 = 0.4191$

Poisson Distribution for acceptance & rejection

- set $n = 0$ & $P = 1$
- Generate random no. $R_{x+1} \sim U[0,1]$ & replace P by $P * (R_{n+1})$
- If $P < e^{-\lambda}$ then accept $N = n$
- Else reject n & increment n by 1 & return to step 2

~~NSPP~~

Q) The no. of customers arriving at a bank b/w 10 am to 1 pm is poisson's distributed with mean = 3. Generate poisson's random variates using random numbers $R_1 = 0.5294$; $R_2 = 0.3578$ and $R_3 = 0.1574$.

$n = 0$ $\qquad P = 1$

Replace P by $P \times R_{n+1}$ $\qquad P = 1 \times 0.5294 = 0.5294$

$$e^{-\lambda} = e^{-3} = 0.04$$

$$0.5294 < 0.04$$

Rejected

$n = 1$

$$P = 0.5294 \times 0.3578 = 0.1894$$

$$0.1894 < 0.04$$

Rejected

$n = 2$

$$P = 0.1894 \times 0.1574 = 0.0298$$

$$0.0298 < 0.04$$

Accepted.

NSPP

→ Let $\lambda^* = \max \lambda(t)$ be the max of the arrival rate function & set $t = 0$ and $i = 1$

→ Generate

S)

Arrivals per period	frequency
0	12
1	10
2	19
3	17
4	10
5	8
6	7
7	5
8	5
9	3
10	3
11	1

Mean $= \Sigma x_i f_i / n$

$= 0 + 10 + 38 + \cancel{36}\ 51 + 40 + 40$

$= \frac{364}{100} = 3.64$

$\Sigma f_i x_i^2 = 2080$

$s^2 = \frac{2080 - 100^* (3.64)^2}{99}$

$= 7.63$

Q. No. of vehicles arriving at an intersection b/w 7.00 am & 7.05 am was monitered for 100 random workdays

Arrivals per period	frequency
0	12
1	10
2	19
3	17
4	10
5	8
6	7
7	5
8	5
9	3
10	3
11	1

H0: the random variable is Poisson distributed

H1: the random variable is not Poisson distributed.

Apply the chi-square test to these data to test if underlying distribution is poisoon. $\alpha = 0.05$

$$P(X) = \frac{e^{-\alpha} \alpha^x}{x!} \quad x = 1, 2$$

$$\bar{X} = \frac{\sum_{j=1}^{n} f_j x_j}{\sum f_j}$$

$$= \frac{10 + 38 + 51 + 40 + 40 + 42 + 35 + 40 + 27 + 30 + 11}{100}$$

$$\lambda = \frac{364}{100} = \cancel{32.33}\ \underline{3.64}$$

$$P(x) = \frac{e^{-3.64}\,(3.64)^x}{x!}$$

$$P(0) = \frac{e^{-3.64}\,(3.64)^0}{0!} = 0.026$$

$$P(1) = \frac{e^{-3.64}\,(3.64)^1}{1!} = 0.096$$

$P(2) = 0.174$

$P(3) = 0.211$

$P(4) = 0.192$

$P(5) = 0.140$

$P(6) = 0.085$

$P(7) = 0.044$

$P(8) = 0.020$

$P(9) = 0.008$

$P(10) = 0.003$

$P(11) = 0.001$

$E(X) = np.$

Arrivals per period	Frequency		Expected		$(O-E)^2/E$
0	12	22	2.6	12.2	7.872 ~~83.985~~ ~~0.803~~
1	10		9.6		~~0.0167~~
2	19		17.4		0.1471
3	17		21.1		0.7967
4	10		19.2		4.4083
5	8		14.0		2.5774
6	7		8.5		0.2647
7	~~15~~	17	4.4	7.6	11.6263
8	5		2.0		
9	3		0.8		
10	3		0.3		
11	1		0.1		

$\chi^2 = 27.6802$

$\chi^2_{\alpha, k-s-1}$ $\chi^2_{0.05,5} = 11.1$

$\downarrow \quad \downarrow$

$7 \quad 1$

$\chi_0^2 > \chi^2_{0.05,5}$

<u>Reject</u> <u>H_0</u>

A NGO collected the records of monthly number of job related injuries in ~~the~~ an industry. The records for the past 100 months are as follows:

Injuries per month	frequency
0	35
1	40
2	13
3	6
4	4
5	1
6	1

Q4) Life tests were performed on a random sample of electronic components at 1.5 times the nominal voltage, & their lifetime in days was recorded.
Apply the chi-square distribution to these data to test the hypothesis that the underlying distribution is exponential (Assume k=8).

H0 : the random variable is exponential distributed
H1 : the random variable is not exponential distributed

$$F(a_i) = 1 - e^{-\lambda a_i}$$

$$\hat{\lambda} = \frac{1}{\bar{X}} = 0.084$$

Let k=8, each interval will have probability p= 0.125

$$ip = 1 - e^{-\lambda a_i}$$

$$\Rightarrow a_i = -\frac{1}{\lambda} \ln(1 - ip)$$

$$a_1 = -\frac{1}{0.084} \ln(1 - 0.125) = 1.590$$

a_1 = 1.590

a_2 = 3.425

a_3 = 5.595

a_4 = 8.252

a_5 = 11.677

a_6 = 16.503

a_7 = 24.75

Class Interval	Observed	P	Expected	$(O-E)^2/E$
[0, 1.590)	19	0.125	6.25	26.01
[1.590, 3.425)	10	0.125	6.25	2.25
[3.425, 5.595)	3	0.125	6.25	0.81
[5.595, 8.252)	6	0.125	6.25	0.01
[8.252, 11.677)	4	0.125	6.25	4.41
[11.677, 16.503)	1	0.125	6.25	4.41
[16.503, 24.755)	4	0.125	6.25	0.8
[24.755, ∞)	6	0.125	6.25	0.01
				39.6

$\chi_0^2 = 39.6 > \chi^2_{0.05, 8-1-1} = 12.6$

Reject H_0.

Q) The time required for 50 different employees to compute & record the number of hours worked during the week with the following results (in mins) given in table below.

1.88, 0.54, 1.90, 0.15, 0.02, 2.81, 1.50, 0.53, 2.62, 2.67, 3.53, 0.53, 1.80, 0.79, 0.21, 0.80, 0.26, 0.63, 0.36, 2.03, 1.42, 1.28, 0.82, 2.16, 0.05, 0.04, 1.49, 0.66, 2.03, 1.00, 0.39, 0.34, 0.01, 0.10, 1.10, 0.24, 0.26, 0.45, 0.17, 4.29, 0.80, 5.50, 4.91, 0.35, 0.36, 0.90, 1.03, 1.73, 0.38, 0.48

Test the hypothesis that these service times are exponentially distributed. $k=6$, $\alpha = 0.05$

HO: the random variable is exponential distributed
H1: the random variable is not exponential distributed

$$F(a_i) = 1 - e^{-\lambda a_i}$$

$$\hat{\lambda} = \frac{1}{\bar{X}} = \frac{1}{1.206} = 0.829$$

$$p = 0.1667$$

$$a_i = -\frac{1}{\lambda} \ln(1 - ip)$$

$$a_1 = -\frac{1}{0.829} \times \ln(1 - 0.1667) = 0.22$$

$a_2 = 0.49$
$a_3 = 0.84$
$a_4 = 1.33$
$a_5 = 2.16$
$a_6 =$

Class Interval	Observed	Expected	$(O-E)^2/E$
[0, 0.22)	8	8.33	0.013
[0.22, 0.49)	11	8.33	0.856
[0.49, 0.84)	9	8.33	0.0539
[0.84, 1.33)	5	8.33	1.331
[1.33, 2.16)	9	8.33	0.0539
[2.16, ∞)	8	8.33	0.013

$\chi_0^2 = 2.3208$

$$\chi_{0.05,\ 6-1-1} = 9.45$$

$$\chi_0^2 \le \chi_{0.05,4}$$

Ho is accepted.

K-S Test for goodness of fit.

- Sample size is small
- No parameters is taken into account.

Q) The highway between Mumbai-Delhi & Calcutta-Delhi has high incidents of accidents along its 100 kms. Public safety officers say that the occurance of these accidents along the highway is randomly (uniformly) distributed, but the news media say otherwise. The Delhi Department of public safety published records for month of June. These records indicated the spot at which 30 accidents involving an injury or death occured as follows: ~~Inside the analysis~~ (the data points represent distance from Mumbai)

88.3, 40.7, 36.3, 27.3, 36.8, 91.7, 67.3, 7.0, 45.2, 23.3, 98.8, 90.1, 17.2, 23.7, 97.4, 32.4, 87.8, 69.8, 62.6, 99.7, 20.6, 73.1, 21.6, 6.0, 45.3, 76.6, 73.2, 27.3, 87.6, 87.2.

use KS Test to determine the goodness of fit for this data with $\alpha = 0.05$

0.883, 0.407, 0.363, 0.273, 0.368, 0.917, 0.673, 0.07, 0.452, 0.233, 0.988, 0.901, 0.172, 0.237, 0.974, 0.324, 0.878, 0.698, 0.626, 0.997, 0.206, 0.731, 0.216, 0.06, 0.453, 0.766, 0.732, 0.273, 0.876, 0.872

i	$R(i)$	i/N	$i/N - R$	$R(i) - \frac{(i-1)}{N}$
1	0.06	0.0333	-	0.06
2	0.07	0.0667	-	0.0367
3	0.172	0.1	-	0.1053
4	0.206	0.1333	-	0.106
5	0.216	0.1667	-	0.0493
6	0.232	0.2	-	0.0653
7	0.237	0.2333	-	0.037
8	0.273	0.2667	-	0.0397
9	0.273	0.3	0.027	0.0063
10	0.324	0.3333	-	0.024
11	0.363	0.3667	0.0037	0.0297
12	0.368	0.4	0.032	0.0013
13	0.407	0.4333	0.0263	0.007
14	0.452	0.4667	0.0147	0.0187
15	0.453	0.5	0.047	-
16	0.626	0.5333	-	0.126
17	0.673	0.5667	-	0.1397
18	0.698	0.6	-	0.1313
19	0.731	0.6333	-	0.131
20	0.732	0.6667	-	0.0987
21	0.766	0.7	-	0.0993
22	0.872	0.7333	-	$\boxed{0.172}$
23	0.876	0.7667	-	0.1427
24	0.878	0.8	-	0.1113
25	0.883	0.8333	-	0.083
26	0.901	0.8667	-	0.0677
27	0.917	0.9	-	0.0503
28	0.974	0.9333	-	0.074
29	0.988	0.9667	-	0.0547
30	0.997	1	0.03	0.0303
			0.047	0.172 (D)

$D_{max} \leq D_{0.05,30} = 0.24$ HO is accepted

$$\mathrm{cov}(x_1, x_2) = \frac{1}{n-1}\left(\sum_{j=1}^{n} x_{1j}\, x_{2j} - n\bar{x}_1 \bar{x}_2\right)$$

$$\hat{\rho} = \frac{\hat{\mathrm{cov}}(x_1, x_2)}{\hat{\sigma}_1 \hat{\sigma}_2}$$

Q) A local home centre specialises in selling designer dirt has streamlined his checkout process. They believe that they are best, highest spending customer expiriencing long service times & are considering adding a separate channel for these preferred customer. Table below gives the list of purchased amounts & service times of 15 customers. Calculate the correlation b/w these data to verify or refute the belief.

Customers	Purchase	Service Time	$X_1 X_2$
1	101	2.2	222.2
2	177	4.4	778.8
3	334	2.1	701.4
4	240	6.9	1656
5	429	5.6	2402.4
6	170	7.3	1241
7	384	8.5	3264
8	222	5.1	1132.2
9	137	4.3	589.1
10	159	7.2	1144.8
11	495	7.6	3762
12	148	5.0	740
13	104	5.3	551.2
14	313	3.5	1095.5
15	341	7.5	2557.5
			21838.1

$\bar{X}_1 = 250.2667$ $\quad \bar{X}_2 = 5.5$

$-n\bar{X}_1 \bar{X}_2$

...0725 $\quad$ -85.076

$$\text{cov}(X_1, X_2) = \frac{1}{n-1}\left(\sum_{j=1}^{n} X_{1j} X_{2j} - n\bar{X}_1\bar{X}_2\right)$$

$$\hat{\rho} = \frac{\hat{\text{cov}}(X_1, X_2)}{\hat{\sigma}_1 \hat{\sigma}_2}$$

Q) A local home centre specialises in selling designer dirt has streamlined his checkout process. They believe that they are best, highest spending customer experiencing long service times & are considering adding a separate channel for these preferred customer. Table below gives the list of purchased amounts & service times of 15 customers. Calculate the correlation b/w these data to verify or refute the belief.

Customers	Purchase	Service Time	$X_1 X_2$
1	101	2.2	222.2
2	177	4.4	778.8
3	334	2.1	701.4
4	240	6.9	1656
5	429	5.6	2402.4
6	170	7.3	1241
7	384	8.5	3264
8	222	5.1	1132.2
9	137	4.3	589.1
10	159	7.2	1144.8
11	495	7.6	3762
12	148	5.0	740
13	104	5.3	551.2
14	313	3.5	1095.5
15	341	7.5	2557.5
	$\bar{X}_1 = 250.2667$	$\bar{X}_2 = 5.5$	21838.1

$-n\bar{X}_1\bar{X}_2$

$$\text{cov}(X_1, X_2) = \frac{1191.0725}{14} = 85.076$$

Inverse - Transform Technique

Exponential: $x_i = \frac{-1}{\lambda} \ln(R_i)$

Uniform: $x_i = a + (b-a)R$

Triangular: $x_i = \begin{cases} \sqrt{2R_i} & 0 \le R_i \le \frac{1}{2} \\ 2 - \sqrt{2(1-R_i)} & \frac{1}{2} \le R_i \le 1 \end{cases}$

Emperical: $X = x_{(i-1)} + a_i \left(R - \frac{(i-1)}{n} \right)$

$$a_i = \frac{x_{(i)} - x_{(i-1)}}{1/n}$$

Exponential:

$E(X) = 1/\lambda$ $\quad V(X) = 1/\lambda^2$

$f(x) = \lambda e^{-\lambda x}$ $\quad F(X) = 1 - e^{-\lambda x}$

Normal:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \quad -\infty < x < \infty$$

$$F(X) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right] dt$$

Standard-Normal:

$\mu = 0 \quad \sigma^2 = 1$

$$Z = \frac{X-\mu}{\sigma} \qquad F(X) = \phi\left(\frac{x-\mu}{\sigma}\right)$$

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

Triangular:

$E(X) = a+b+c/3.$

$$f(x) = \begin{cases} \frac{2(x-a)}{(b-a)(c-a)} & a \le x \le b \\ \frac{2(c-x)}{(c-b)(c-a)} & b < x \le c \\ 0 & \text{elsewhere} \end{cases}$$

$$F(X) = \begin{cases} 0, & x \le a \\ \frac{(x-a)^2}{(b-a)(c-a)}, & a < x \le b \\ 1 - \frac{(c-x)^2}{(c-b)(c-a)}, & b < x \le c \\ 1 & x > c \end{cases}$$

classmate
Date
Page

Statistics:

$P(a < X \leq b) = F(b) - F(a)$

CDF: $F(x) = \sum_{x_i \leq x} p(x_i)$

$F(x) = \int_{-\infty}^{x} f(t)\, dt$

Mean: $E(x) = \sum_{\text{all } i} x_i\, p(x_i)$

$E(x) = \int_{-\infty}^{\infty} x f(x)\, dx$

variance: $V(X) = E[(X - E[X])^2]$

$V(X) = E(X^2) - [E(X)]^2$

Bernoulli: $E(X) = p$ $\quad V(X) = pq$

Binomial: $E(x) = np$ $\quad V(X) = npq$

$p(x) = {}^nC_x\, p^x q^{n-x}$ $\quad q = 1 - p.$

Geometric: $E(x) = 1/p$ $\quad V(X) = q/p^2$

$P(x) = q^{x-1} p$ $\quad F(x) = 1 - q^x$

Poisson: $E(X) = V(X) = \alpha$

$p(x) = \dfrac{e^{-\alpha} \alpha^x}{x!}$ $\quad F(x) = \sum_{i=0}^{x} \dfrac{e^{-\alpha} \alpha^i}{i!}$

Uniform: $E(X) = a+b/2$ $\quad V(X) = (b-a)^2/12$

$f(x) = \dfrac{1}{b-a} \quad a \leq x \leq b.$

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

→ Add the above value, to last ~~observed~~ expected value of chi-square.

* Repeat chi-square process

Autocorrelation:

$$Z_0 = \frac{\hat{\rho}_{im}}{\hat{\sigma}_{im}} \qquad i + (M+1)m \leq N$$

i = pattern start index
m = differing pattern value
N = no. of observations

$$\hat{\rho}_{im} = \frac{1}{M+1}\left[\sum_{k=0}^{M} R_{i+km} \cdot R_{i+(k+1)m}\right] - 0.25$$

$$\hat{\sigma}_{\rho im} = \frac{\sqrt{13M+7}}{12(M+1)}$$

Gaps:

→ For every digit, calculate no. of gaps & length of gaps

Total Number of Gaps (n) = No. of digits − Number of distinct digits.

Gap length	Frequency	Relative Frequency	CF $F(x)$	S_N Theoretical CF $1-(0.9)^{x+1}$	$\lvert F(x) - S_N \rvert$ ↓ max t_{calc}

t_α is removed from KS table.

$t_{calc} < t_\alpha \Rightarrow H_0$ is accepted.

Runs above & below mean:

$n1$ = +ve count

$n2$ = -ve count

$N = n1 + n2$

b = no. of runs

$Z_0 = \frac{b - \mu_b}{\sigma_b}$

$\mu_b = \left(\frac{2n_1 \cdot n_2}{N}\right) + \frac{1}{2}$

$\sigma_b^2 = \frac{2n_1 \cdot n_2 (2n_1 n_2 - N)}{N^2(N-1)}$

Length of runs:

→ +, −

→ sequence of runs

→ Run length

Observed

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

$E(Y_i) = \frac{2}{N!}$ for $i = N-1$

Runs up & down

→ $E(Y_i) = \frac{N w_i}{E(I)}$

$w_i = \left(\frac{n1}{N}\right)^i \left(\frac{n2}{N}\right) + \left(\frac{n1}{N}\right)\left(\frac{n2}{N}\right)^i$

$E(I) = \frac{n1}{n2} + \frac{n2}{n1}$

Runs above & below

→ ~~expected total no. of runs~~ mean

$\mu a = \frac{2N-1}{3}$; Runs up & down

$E(A) = \frac{N}{E(I)}$; Runs above & below

→ Expected no. of runs length

$\mu a - \sum_{i=1}^{m} E(Y_i)$ m: max length of observed r

~~Runs~~ above & below mean:

$n1$ = +ve count

$n2$ = -ve count

$N = n1 + n2$

b = no. of runs

$Z_0 = \frac{b - \mu_b}{\sigma_b}$

$\mu_b = \left(\frac{2n_1 \cdot n_2}{N}\right) + \frac{1}{2}$

$\sigma_b^2 = \frac{2n_1 \cdot n_2 (2n_1 n_2 - N)}{N^2(N-1)}$

Length of runs:

- +, -
- sequence of runs
- Run length / Observed

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

$$E(Y_i) = \frac{2}{N!} \quad \text{for } i = N-1$$

Runs up & down

$$E(Y_i) = \frac{N w_i}{E(I)}$$

$$w_i = \left(\frac{n1}{N}\right)^i \left(\frac{n2}{N}\right) + \left(\frac{n1}{N}\right)\left(\frac{n2}{N}\right)^i$$

$$E(I) = \frac{n1}{n2} + \frac{n2}{n1}$$

Runs above & below

→ ~~Expected total no. of runs~~ mean

$\mu a = \frac{2N-1}{3}$; Runs up & down

$E(A) = \frac{N}{E(I)}$; Runs above & below

→ Expected no. of runs length

$\mu a - \sum_{i=1}^{m} E(Y_i)$

m: max length of observed runs.

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

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

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

$H_0: R_i \sim$ independently

$H_1: R_i \nsim$ independently

For autocorrelation: $H_0: \rho_{im} = 0$, if numbers are independent

$H_1: \rho_{im} \neq 0$, if numbers are dependent

Kolmogorov-Smirnov Test

N: no. of observations

→ Arrange observations ascending

i

$R(i)$

i/N

$i/N - R(i)$ → max (D^+)

$R(i) - \frac{i-1}{N}$ → max (D^-)

} max (D)

If getting -ve values use '-'

$D \leq D_\alpha$, null hypothesis is accepted.

Chi-Square Test

$E_i = \frac{N}{n}$ usually assume > 5.

N: No. of observations

n: No. of classes

$$\chi_0^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

Ensure all observed & expected values are ≥5.

DOF = n-1

↳ no. of classes after adjustment

$\chi_0^2 \leq \chi^2_{\alpha, dof}$, Ho is accepted.

Runs:

Up & Down:

No. of runs → a

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

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

$-Z_{\alpha/2} \leq Z_0 \leq Z_{\alpha/2}$, Ho is accepted.

Newspapers:

profit = (revenue from sales) − (cost of newspaper) − (lost profit from excess demand) + (salvage from sale of scrap papers)

M–N Inventory:

M → max inventory level

N → Review period

Days, RD for demand, demand, Start Inventory, ending Inventory, shortage, order quantity, RD for lead time, Days until Order arrives.

LCM:

$$x_{i+1} = (ax_i + c) \bmod m$$

$c \neq 0$: mixed congruential method

$c = 0$: multiplicative " "

Case 1: $m = 2^b$

$c \neq 0$

$P = m$

c is relatively prime to m

$a = 1 + 4k$

$$\text{Diff} = \sum \frac{|\text{num} - \text{current}|}{m}$$

$$\text{Density} = \frac{\text{Diff}}{\text{period}}$$

Case 2: $m = 2^b$

$c = 0$

$P = m/4$

X_0 is odd

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

Case 3: m is prime

$c = 0$

$P = m - 1$

$a^k - 1$ is divisible by m

$k = m - 1$

classmate

Date
Page

Lead-Time: (As many demand raws as lead time) (Lead-Demand Time)

Reliability: (Cumulative lifetime)

when replacement multiply by 3.

Cost of bearings = Total No. of bearings × Cost per bearing

Cost of delay time = Total delay time × Downtime/min

Cost of downtime during repairs = No. of bearings × Bearing Time × Downtime/min

Cost of repairpersons = No. of bearings × Bearing Time × Cost of repair/min

Random Normal Numbers:

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

$$X = Z\sigma_X$$

$$Y = Z\sigma_y$$

Grocery problem: (Service Begins, wait time, Service ends, Time in system, Idle time of server)

Avg. waiting time = $\frac{\text{Total time customers wait in queue}}{\text{Total number of Customers}}$

Probability (wait) = $\frac{\text{Numbers of customers who wait}}{\text{Total number of customers}}$

Probability of Idle server = $\frac{\text{Total idle time of server}}{\text{Total run time of simulation}}$

Average service time = $\frac{\text{Total service Time}}{\text{Total number of Customers}}$

Average time between arrivals = $\frac{\text{Sum of all times b/w arrival}}{\text{No. of arrivals} - 1}$

Average waiting time of those who wait = $\frac{\text{Total time customers wait in queue}}{\text{Total number of customers that wait}}$

Average Time customer spends in system = $\frac{\text{Total system Time}}{\text{Total number of customers}}$