for optimum oil yield from palm kernel. ... are required

Basic terms in STATISTICS

(1) Descriptive Stat: A type of statistics that describe and summarize data.

(2) Inferential statistics: A part of statistics that is concerned with drawing conclusion from data value.

(3) Population: A collection of element of interest.

(4) Sample: Sample is a subgroup of population that is to studied

(5) Qualitative: They are measurements with no natural numerical state but consist of label and non numerical characteristics.

8) Quantitative: These are measurement that arise from natural numerical scale.

7) Probability model: These are numbers that represents a probability as a whole.

Theorem Of Probability

Question 1: Show that $P(A') = 1 - P(A)$

Soln

recall $A \cup A^c = S$ ---(i)

$A \cap A^c = \phi$ ---(ii)

from eqn (i)

$A \cup A^c = S$

$P(A \cup A^c) = P(S)$ ... (iii)

$P(A) + P(A^c) - P(A \cap A^c) = P(S)$

from eqn (iii)

$P(A \cap A^c) = P(S)$ $\left(P(m) = \frac{n(m)}{n(S)} \text{ that's what we used} \right)$

$\frac{n(A \cup A^c)}{n(S)} = \frac{n(S)}{n(S)}$

$\frac{n(A \cup A^c)}{n(S)} = 1$

$n(A \cup A^c) = n(S)$

reverse case

$P(A \cup A^c) = 1$

$P(A) + P(A^c) - P(A \cap A^c) = 1$

recall $(A \cap A' = \emptyset)$

$P(A) + P(A^c) - P(\phi) = 1$

so $P(A^c) = 1 - P(A)$ //

Question 2: Show that

$P(\emptyset) = 0$

Soln

recall $S \cup \phi = S$

$P(S \cup \phi) = P(S)$

$P(S) + P(\phi) - \cancel{P(S \cap \phi)} = P(S)$

$P(\phi) - \cancel{P(S \cap \phi)} = P(S) - P(S)$

$P(\phi) - \cancel{P(S \cap \phi)} = 0$

$P(\phi) = \cancel{P(S \cap \phi)} = 0$

$P(\phi) = 0$ //

Question 3: Show that

$\emptyset \leq P(A) \leq 1$

Soln

Assuming ϕ depend on A, and A depend on Sample space

$\phi \subset A \subset S$

$P(\phi) \leq P(A) \leq P(S)$

$\emptyset \leq A \leq 1$

$\emptyset \leq P(A) \leq 1$

... and 150°C with constant velocity of 1m/s on the quantity of palm kernel oil produced or expelled.

Theory Of Probability

* $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

* If A and B occur in a sample space and mutually exclusive, $P(A \cap B) = 0$, $P(A \cup B) = P(A) + P(B)$

* If A and B are independent of each other.

$P(A \cap B) = P(A) \cdot P(B)$

$P(A \cup B) = P(A) + P(B) \cdot P(A \cap B)$

* Conditional Probability

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

$$P(B/A) = \frac{P(A \cap B)}{P(A)}$$

$$P(A'/B') = \frac{P(A' \cap B')}{P(B')} = \frac{P(A \cap B)'}{P(B')}$$

$$\frac{P(A \cap B)'}{P(B')} = \frac{1 - P(A \cup B)}{P(B')}$$

$$= \frac{1 - P(A) + P(B) - P(A \cap B)}{P(B')}$$

Question ? A box of 100 items, 10 with type A defect, 5 items with type B defect. 2 items with both type A and B deffect. Find the prob that (i) An item drawn is a type B defect under the condition that it is a type A defect.

(ii) An item drawn has no type B deffect under the condition that it has no type A deffect.

Soln

$n(S) = 100$, $n(A) = 10$, $n(B) = 5$

$n(A \cap B) = 2$.

$$P(A) = \frac{n(A)}{n(S)} = \frac{10}{100} = 0.1$$

$$P(B) = \frac{n(A)}{n(S)} = \frac{5}{100} = 0.05$$

$$P(A \cap B) = \frac{n(A \cap B)}{n(S)} = \frac{2}{100} = 0.02$$

(i) $$P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{0.02}{0.1}$$

$$= 0.2$$

(ii) $$P(B'/A') = \frac{1 - P(A) + P(B) - P(A \cap B)}{P(A')}$$

$$= \frac{1 - P(A) + P(B) - P(A \cap B)}{1 - P(A)}$$

$$= \frac{1 - 0.1 + 0.05 - 0.02}{} = \cancel{1.035}\ 0.964$$

Question : If the Probability of A = 0.4, P(B) = 0.7. If the probability of at least One of A and B is 0.8. Find the probability of only one of A and B.

Soln

$P(A) = 0.4, \quad P(B) = 0.7$

$P(A \cap B) = 0.8, \quad P(A \cup B) = ?$

Soln

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$= 0.4 + 0.7 - 0.8$

$= 0.3$

Question! A personel department of Company shows the following analysis

Age	Bachelor degree	Master degree	Total
Under 30	90	10	100
30–40	20	30	50
over 40	40	10	50
Total	150	50	200

If one engineer is selected at random from the Company, find

(i) P(only bachelor degree)

(ii) P(A master degree given that he is over 40)

(iii) P(he is under 30 given that he has only bachelor degree)

Soln

$P(\text{only bachelor}) = \frac{150}{200} = \frac{3}{4}$

(ii) $P(\text{master degree over 40}) = \frac{10}{50} = 0.2$

(iii) $P\left(\frac{\text{under 30}}{B}\right) = \frac{90}{150}$

Question : A Shop is supplied with goods manufactured by three factories whose relative [illegible] are

Factory 1 = 50%, Factory 2 : 30%

Factory 3 = 20%.

The Percentage of deffective goods manufactured by these factories, 2%, 3% and 5% respectively

What is the probability that an article purchased at random at this shop will turn out to be non-defective item.

Soln

$P(\text{Factory 1}) = \frac{50}{100} = 0.50$

$P(\text{factory 2}) = \frac{30}{100} = 0.30$

$P(\text{factory 3}) = \frac{20}{100} = 0.20$

$P\left(\frac{D}{F_1}\right) = \frac{2}{100} = 0.02$

$P\left(\frac{D}{F_2}\right) = \frac{3}{100} = 0.03$

$P\left(\frac{D}{F_3}\right) = \frac{5}{100} = 0.05$

$P\left(\frac{\text{Non D.}}{F_1}\right) = \frac{P(F_1) \cdot P\left(\frac{D}{F_1}\right)}{P(F_1) \cdot P\left(\frac{D}{F_1}\right) + P(F_2) \cdot P\left(\frac{D}{F_2}\right)}$

$= 0.34$

(ii) $P\left(\frac{\text{non D}}{F_2}\right) = \frac{P(F_2) \cdot P\left(\frac{D}{F_2}\right)}{P(F_1) \cdot P\left(\frac{D}{F_1}\right) + P(F_2) \cdot P\left(\frac{D}{F_2}\right)}$

CSE 331

Correlation & Regression Analysis

$$\hat{y} = \beta_1 x + \beta_0 + \varepsilon \text{ (error)}$$

where

$$\beta_1 = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

$$\beta_0 = \bar{y} - \beta_1 \bar{x}$$

~~Second Method~~

C.O.D Coefficient of Determinant

$$COD(r^2) = \frac{\text{Sum of Square of regression}}{\text{Sum of square of total}}$$

$$= \frac{\sum(\hat{y}-\bar{y})^2}{\sum(y-\bar{y})^2}$$

$$C.O.C\ (r) = \pm\sqrt{C.O.D}$$

$$\underset{\text{Sum Square total}}{SST} = \underset{\text{sum square regression}}{SSR} + \underset{\text{sum square error}}{SSE}$$

Question

Find the distribution function which correspond to P.D.F given $f(x) = \begin{cases} 3e^{-3x} & \text{for } x > 0 \\ 0 & \text{else} \end{cases}$

find $P(0.5 \le x \le 1)$

Soln

$x > 0$

$$\int_0^\infty 3e^{-3x}\,dx = 3\left[\frac{e^{-3(\infty)}}{-3} - \frac{e^{-3(0)}}{-3}\right]$$

$$3\left[0 - \frac{1}{-3}\right]$$

$$= -3\left(\frac{-1}{3}\right) = \frac{3}{3} = 1$$

(ii) $P(0.5 \le x \le 1)$

$$\int_{0.5}^{1} 3e^{-3x}\,dx$$

$$= 3\left[\frac{e^{-3(1)}}{-3} - \frac{e^{-3(0.5)}}{-3}\right]$$

$$= 0.173$$

6b

Age	35	40	38	44	67	64	59	69	25	50
Bp	112	128	130	138	158	162	140	175	125	142

(i) Derive the regression model for Patient age and blood pressure

Soln

y is BP cos it depends on Age (x).

x	y	xy	x^2	$\hat{y}$	$\hat{y}-\bar{y}$	$(\hat{y}-\bar{y})^2$	$y-\bar{y}$
35	112	3920	1225	168.904	27.904	778.633	-29
40	128	5120	1600	159.009	18.009	524.324	-13
38	130	4940	1444	162.967	21.964	482.549	-11
44	138	6072	1936	159.093	18.095	327.429	-3
67	158	10586	4489	105.576	-35.424	1254.39	17
64	162	10368	4096	111.513	-25.487	869.483	21
59	140	8260	3481	121.408	-19.593	383.586	-1
69	175	12075	4761	101.618	-39.38	155.942	34
25	125	3125	625	188.694	47.694	229.718	-16
50	142	710	2500	139.219	-1.781	8.172	1
491	1410	71563	26157		6202	6202.995	

hence it is a viscous semi-solid, even at tropical ambient and a solid fat in temperate climates

$$\beta_1 = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

$$= \frac{10(65176) - (491 \times 1410)}{10(2615) - (491)^2}$$

$\bar{x} = 49.1$

$\beta_1 = -1.9796$

$\bar{y} = \frac{1410}{10} = 141$

7

$\beta_0 = \bar{y} - \beta\bar{x}$

$\beta_0 = 141 - (-1.979 \times 49.1)$

$\beta_0 = 238.169$ //

$y = \beta_1 x + \beta_0 + \varepsilon$

$= (-1.979)x + 238.169$

$y = -1.979x + 238.16$

so we use this to get our y at different values of x.

$$C.O.D = \frac{\Sigma(\hat{y}-\bar{y})^2}{\Sigma(y-\bar{y})^2} = \frac{6202.995}{3284}$$

$C.O.D(r^2) = \cancel{1.889} = 1.889$

$(C.O.C) = \text{sign}\sqrt{COD}$

$= \text{sign}\sqrt{1.889}$

$= 1.374$

(iii) $\hat{y} = -1.979x + 238.169$

when $x = 25$

$\hat{y} = -1.979(25) + 238.169$

$= 188.694$

when $x = 79$

$\hat{y} = -1.979(79) + 238.169$

$= 81.828$

Class

Poisson Distribution

A random variable x is said to have poisson distribution if the PMF of x is

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

where $x = 0, 1, 2, \ldots$

When the number of trials (n) is very large and the probability of success is comparably small the binomial distribution will not be very suitable model for the random experiment with repeated trials.

Note that

Mean $= (\lambda)$

Standard deviation $\sigma = \sqrt{np}$

Variance $= np$

Example 1: One out of a thousand people reacted to a newly manufactured vaccine against ebola virus disease. If 3000 people were treated with this vaccine. Find the probability that

(a) At most two people reacted to the vaccine

(b) At least 3 people reacted

(c) None reacted

(d) Not less than one person reacted to the vaccine pressed by twin-screw expellers or hydraulically to yield red crude oil. (Source?)

soln

$n = 3000$

$p = 1/1000 = 0.001$

$\lambda = np = 3000 \times 0.001$

$\lambda = 3$

(a) $P(x \leq 2) = P(x=0) + P(x=1) + P(x=2)$

0, 1, 2

$$= \frac{e^{-3} \cdot 3^0}{0!} + \frac{e^{-3} \cdot 3^1}{1!} + \frac{e^{-3} \cdot 3^2}{2!}$$

$P(x \leq 2) = 0.423$ //

(b) $P(r \geq 3) = 1 - P(r \leq 2)$

$= 1 - 0.423$

$= 0.577$ //

(c) $1 - P(x=0)$ (means not less than one reacted)