

INSTITUTE OF ACTUARIES OF INDIA

Subject CS1A – Actuarial Statistics (Paper A)

November 2024 Examination

INDICATIVE SOLUTION

Introduction:

The indicative solution has been written by the Examiners with the aim of helping candidates. The solutions given are only indicative. It is realized that there could be other points as valid answers and examiners have given credit for any alternative approach or interpretation which they consider to be reasonable.

Solution 1: Correct Answer is **Option B.** [2]

We know that sample correlation coefficient = $S_{xy} / \sqrt{S_{xx} * S_{yy}}$.

Beta coefficient in the regression line is given by S_{xy} / S_{xx} .

So, the sign (positive or negative) of both the correlation coefficient and beta coefficient is dependent on the sign of S_{xy} . It follows that if beta coefficient in the regression line is positive then correlation coefficient is also positive.

Looking at the each of the statements –

- A. β_1 is positive, so Y and X_1 cannot be negatively correlated.
- B. β_2 is positive, so Y and X_2 can be positively correlated.
- C. β_2 is positive, so Y and X_2 cannot be negatively correlated.
- D. β_1 is positive (non-zero), so Y and X_1 cannot be un-correlated.

Considering the above, only Option B is true.

Solution 2: Correct Answer is **Option C.** [2]

In order to test the population correlation coefficient, null hypothesis and alternative hypothesis are defined as under –

$H_0: \rho = 0$ (in other words, X_1 and X_2 are uncorrelated).

$H_1: \rho \neq 0$ (in other words, X_1 and X_2 are not uncorrelated).

The value of the test statistic under the null hypothesis
 $= r * \sqrt{(n - 2) / (1 - r^2)}$
 $= -0.1599543 * \sqrt{(7 - 2) / (1 - (-0.1599543)^2)}$
 $= -0.36233$.

Solution 3: Correct Answer is **Option D.** [2]

Under H_0 , the test statistic has a students' t-distribution with $n - 2$ degrees of freedom.

For a t_5 distribution, the critical value at 5% level of significance (considering a two-sided hypothesis) is 2.571.

As the value of test statistic $-0.36233 < 2.571$, we do not have sufficient evidence to reject the null hypothesis. Hence, we conclude that X_1 and X_2 are un-correlated. In other words, we may put on weight or lose weight irrespective of age.

Solution 4: Correct Answer is **Option A.** [2]

$$r_{\text{SPEARMAN}} = 1 - (6 \sum_1^n d^2) / (n * (n^2 - 1))$$

$$\begin{aligned} \text{As } r_{\text{SPEARMAN}} &= 0, \\ 0 &= 1 - (6 \sum_1^7 d^2) / (7 * (49 - 1)) \\ 1 &= (6 \sum_1^7 d^2) / 336 \\ 6 \sum_1^7 d^2 &= 336 \quad \sum_1^7 d^2 = 336 / 6 = 56. \end{aligned}$$

Solution 5: Correct Answer is **Option C.** [2]

$$r_{\text{KENDALL}} = (n_c - n_d) / (n_c + n_d)$$

$(n_c + n_d)$ represents the total number of pairs.

It can be also calculated as $n * (n - 1) / 2$.

$$\begin{aligned} \text{As } n = 7, \text{ the total number of pairs} \\ &= 7 * (7 - 1) / 2 \\ &= 7 * 3 \\ &= 21. \end{aligned}$$

As 21 is an odd number, the number of concordant pairs (n_c) and number of discordant pairs (n_d) cannot be equal. Hence, the value of the numerator ($n_c - n_d$) will be a non-zero number as $n_c \neq n_d$. Consequently, the value of r_{KENDALL} cannot be equal to 0.

Solution 6: Correct Answer is **Option B.** [2]

The paediatrician believes that generally infant height lies in the interval of mean height plus / minus 4 inches.

This can be mathematically represented as

$$P(\text{Mean}(X) - 4 < X < \text{Mean}(X) + 4)$$

X follows a normal distribution and $Y = e^X \sim \log N(\mu = 25, \sigma^2 = 36)$.

Hence, $X \sim N(\mu = 25, \sigma^2 = 36)$.

$$\begin{aligned} \text{So, } P(\text{Mean}(X) - 4 < X < \text{Mean}(X) + 4) \\ &= P((25 - 4) < X < (25 + 4)) \\ &= P(21 < X < 29). \end{aligned}$$

Statements C and D can be outrightly discarded as we are finding probability in terms of X and not in terms of Y.

Let us consider the other two statements –

Statement A –

$$P(|X - 25| < 4) = P(-4 < (X - 25) < 4) = P(21 < X < 29)$$

This does not match with the required probability and hence this statement is incorrect.

Statement B –

$$P(|X - 25| < 4) = P(-4 < (X - 25) < 4) = P(21 < X < 29)$$

This matches with the required probability and hence this statement is the correct answer.

Solution 7: Correct Answer is **Option A.** [2]

$$\begin{aligned} \text{Required probability} \\ &= P(21 < X < 29) \\ &= P(X < 29) - P(X < 21) \\ &= P(Z < (29-25)/6) - P(Z < (21-25)/6) \\ &= P(Z < 0.67) - P(Z < -0.67) \\ &= P(Z < 0.67) - (1 - P(Z < 0.67)) \end{aligned}$$

$$= 0.4952$$

$$= 49.52\%.$$

Solution 8: Correct Answer is **Option D.** [2]

$$E(Y) = \exp(25 + 0.5 * 36) = \exp(43) = 4.73 * 10^{18}$$

$$E(X) = \mu = 25$$

Let us consider each of the given statements.

- A. $E[Y] = E[X]$. This is not true as $E[Y] \neq E[X]$ as seen above.
- B. $E[Y] < E[X]$. This is not true as $E[Y] > E[X]$ as seen above.
- C. $E[Y] = \exp(E[X])$. This is not true as $\exp(25) = 7.2 * 10^{10} \neq E[Y]$
- D. $E[Y] = E[\exp(X)]$. This statement seems to be true. As $Y = \exp(X)$, $E[\exp(X)] = E[Y]$.

Solution 9: Correct Answer is **Option C.** [2]

For a normal distribution,

$$M_x(t)$$

$$= \exp(\mu t + \frac{1}{2} \sigma^2 * t^2)$$

$$= \exp(25t + \frac{1}{2} * 36 * t^2)$$

$$= \exp(25t + 18t^2).$$

Solution 10: Correct Answer is **Option D.** [2]

Let us follow the given steps:

Step 1: Random value u from $U(0, 1)$ is 0.40

Step 2: Not applicable as $u < 0.5$

Step 3: We have to find z such that $P(Z \leq z) = 1 - u$

$$P(Z \leq z) = 1 - 0.40$$

$$P(Z \leq z) = 0.60$$

This can be also expressed as

$$P(Z > z) = 0.40$$

From tables, $z = 0.2533$. Simulated z -value is -0.2533 .

Step 4: Simulated x value = $25 - 0.2533 * 6 = 23.4802$.

Solution 11: Correct Answer is **Option A.** [2]

$$S_{xy}$$

$$= (\sum_1^6 xy - n(x_{\text{bar}})(y_{\text{bar}}))$$

$$= (1062 - 6 * 68/6 * 85/6)$$

$$= 98.67$$

$$S_{xx}$$

$$= (\sum_1^6 x^2 - n(x_{\text{bar}})^2)$$

$$= (842 - 6 * (68/6)^2)$$

$$= 71.33$$

$$\begin{aligned}\beta &= S_{xy} / S_{xx} \\ &= 16.44 / 11.89 \\ &= 1.3832.\end{aligned}$$

Solution 12: Correct Answer is **Option D.**

[2]

$$\begin{aligned}\alpha &= \bar{y} - \bar{x} * \beta \\ &= (85/6) - (68/6) * 1.3832 \\ &= -1.5096\end{aligned}$$

$$\begin{aligned}\hat{y} \text{ (i.e. number of trees expected to fall in one day of the cyclone Pawan)} &= -1.5096 + 1.3832 * (15) \\ &= 19.2384\end{aligned}$$

$$\begin{aligned}\text{So, total number of trees expected to fall in the next three days of cyclone Pawan} &= 19.2384 * 3 \\ &= 57.7152 \\ &= 58 \text{ (rounded off to the next integer as instructed)}\end{aligned}$$

Solution 13: Correct Answer is **Option C.**

[2]

$$\begin{aligned}S_{yy} &= (\sum_1^6 y^2 - n(\bar{y})^2) \\ &= (1463 - 6 * (85/6)^2) \\ &= 258.83\end{aligned}$$

$$\begin{aligned}\text{Residual sum of squares} &= S_{yy} - S_{xy}^2 / S_{xx} \\ &= 258.33 - 98.67^2 / 71.33 \\ &= 122.36 \\ &= 122 \text{ (approximately).}\end{aligned}$$

Solution 14: Correct Answer is **Option B.**

[2]

$$Y = \sigma * \exp(\lambda * X + e)$$

$$\ln Y = \ln \sigma + \lambda * X + e$$

$$\begin{aligned}S_{x \ln y} &= (\sum_1^6 x_i \ln y_i - n(\bar{x})(\bar{\ln y})) \dots \dots \text{ where } \bar{\ln y} = (\sum \ln y_i) / 6 \\ &= (178.519074 - 6 * 68/6 * 15.039379 / 6) \\ &= 8.0728\end{aligned}$$

$$\begin{aligned}\lambda &= S_{x \ln y} / S_{xx} \\ &= 8.0728 / 71.33 \\ &= 0.113170.\end{aligned}$$

Solution 15: Correct Answer is **Option B.**

[2]

$$\begin{aligned}\ln \sigma &= \bar{\ln y} - \lambda * \bar{x}\end{aligned}$$

$$= 15.039379/6 - 0.113170 * 68/6$$

$$= 1.22397$$

$$\sigma$$

$$= \exp(1.22397)$$

$$= 3.400661.$$

Solution 16: Correct Answer is **Option A.** [2]

$$E(X_i)$$

$$= \alpha / \beta$$

$$= 1/2 / (1/10)$$

$$= 5$$

$$\text{Var}(X_i)$$

$$= \alpha / \beta^2$$

$$= 1/2 / (1/10)^2$$

$$= 50.$$

Solution 17: Correct Answer is **Option B.** [2]

$$E(N)$$

$$= n * p$$

$$= (10 * 100) * 0.005$$

$$= 5$$

$$\text{Var}(N)$$

$$= n * p * q$$

$$= (10 * 100) * 0.005 * 0.995$$

$$= 4.975$$

$$\text{SD}(N)$$

$$= \text{sqrt}(4.975)$$

$$= 2.230.$$

Solution 18: Correct Answer is **Option C.** [2]

$$E(Y)$$

$$= E (E(Y|N))$$

$$= E (E(X_1 + X_2 + \dots + X_N))$$

$$= E (N * E(X_i))$$

$$= E (N * 5)$$

$$= 5 * E(N)$$

$$= 5 * 5$$

$$= 25 \text{ lakhs.}$$

Solution 19: Correct Answer is **Option D.** [2]

The correct expression for unconditional variance of Y is:

$$\text{Var}(Y) = E[\text{Var}(Y|N)] + \text{Var}[E(Y|N)].$$

Solution 20: Correct Answer is **Option B.** [2]

$$\text{Var}(Y)$$

$$= E[\text{Var}(X_1 + X_2 + \dots + X_N)] + \text{Var}[E(X_1 + X_2 + \dots + X_N)]$$

$$\begin{aligned}
&= E[N * \text{Var}(X_i)] + \text{Var} [N * E(X_i)] \dots\dots\dots \text{As } X_i\text{s are independent} \\
&= E[N * 50] + \text{Var} [N * 5] \\
&= 50 * E[N] + 5^2 * \text{Var}[N] \\
&= 50 * 5 + 25 * 4.975 \\
&= 374.375
\end{aligned}$$

$$\begin{aligned}
&\text{SD}(Y) \\
&= \text{sqrt}(374.375) \\
&= 19.35.
\end{aligned}$$

Solution 21: Correct Answer is **Option B.** [2]

The 99% two-sided confidence intervals for coefficients β_1 , β_2 and β_3 do not contain the value 0 and hence the explanatory variables corresponding to these coefficients viz. X_1 , X_2 and X_3 can be considered to be significant.

The 99% two-sided confidence interval for coefficient β_4 contains the value 0 and hence the explanatory variable X_4 is considered to be insignificant.

Solution 22: Correct Answer is **Option A.** [2]

$$\begin{aligned}
&\text{SS}_{\text{REG}} \\
&= 1602.748925 * 4 \\
&= 6410.9957.
\end{aligned}$$

$$\begin{aligned}
&\text{SS}_{\text{TOT}} \\
&= 6410.9957 + 2136.9986 \\
&= 8547.9943
\end{aligned}$$

$$\begin{aligned}
&R^2 \\
&= \text{SS}_{\text{REG}} / \text{SS}_{\text{TOT}} \\
&= 6410.9957 / 8547.9943 \\
&= 75\%.
\end{aligned}$$

Solution 23: Correct Answer is **Option D.** [2]

$$\begin{aligned}
&\text{Adjusted } R^2 \\
&= 1 - (n - 1) / (n - k - 1) * (1 - R^2) \\
&= 1 - (25 - 1) / (25 - 4 - 1) * (1 - 75\%) \\
&= 1 - 24/20 * 25\% \\
&= 70\%.
\end{aligned}$$

Solution 24: Correct Answer is **Option B.** [2]

$$\begin{aligned}
&\text{Value of ANOVA Test Statistic} \\
&= \text{regression mean square} / \text{residual mean square} \\
&= 1602.748925 / (2136.9986 / (25 - 4 - 1)) \\
&= 15
\end{aligned}$$

Critical value of $F_{k,(n-k-1)} = F_{4,20}$ at 5% level of significance is 2.866.

The value of the statistic is much greater than the critical value. Hence we have sufficient evidence to reject the null hypothesis.

H_0 states all $\beta_i = 0$ (for $i = 1,2,3,4$). H_1 states that at least one $\beta_i \neq 0$. Hence Option B represents the correct answer. At least one of the explanatory variables has significant impact on the customer value.

Solution 25: Correct Answer is **Option C.** [2]

Selection of the optimal set of explanatory variables (using either forward selection or backward selection) should be based on Adjusted R^2 .

Model	Adjusted R^2	Decision
$X_1 + X_2 + X_3 + X_4$	70%	
$X_1 + X_2 + X_3$	73%	Adjusted R^2 is increasing so X_4 can be dropped.
$X_1 + X_2$	70%	Adjusted R^2 is reducing so X_3 need not be dropped.
X_1	68%	Adjusted R^2 is reducing so X_2 need not be dropped.

So, the optimal set of explanatory variables will be X_1 , X_2 and X_3 . This is consistent with the solution obtained in question 21.

Solution 26: Correct Answer is **Option D.** [2]

Type I Error: Reject H_0 when it is true.

Type II Error: Fail to reject H_0 when it is false.

Solution 27: Correct Answer is **Option A.** [2]

$$\begin{aligned}
 &\text{Probability of Type I Error} \\
 &= P(X > 93.50) \text{ under } H_0: (\lambda = 50) \\
 &= \int_{93.50}^{\infty} (3\lambda^3 (\lambda + x)^{-4}) dx \\
 &= -3\lambda^3 / 3 * (\text{from } 93.50 \text{ to } \infty (\lambda + x)^{-3}) \\
 &= -50^3 * (\text{from } 93.50 \text{ to } \infty (50 + x)^{-3}) \\
 &= 0.0423 \\
 &= 4.23\%
 \end{aligned}$$

Solution 28: Correct Answer is **Option B.** [2]

$$\begin{aligned}
 &\text{Probability of Type II Error} \\
 &= P(X \leq 93.50) \text{ under } H_1: (\lambda = 60) \\
 &= \int_0^{93.50} (3\lambda^3 (\lambda + x)^{-4}) dx \\
 &= -3\lambda^3 / 3 * (\text{from } 0 \text{ to } 93.50 (\lambda + x)^{-3}) \\
 &= -60^3 * (\text{from } 0 \text{ to } 93.50 (60 + x)^{-3}) \\
 &= 0.9403 \\
 &= 94.03\%
 \end{aligned}$$

Solution 29: Correct Answer is **Option C.** [2]

Power of Test = 1 – Probability of committing Type II Error
 Size of Test = Probability of committing Type I Error

Solution 30: Correct Answer is **Option C.** [2]

Sensitivity of Test = Power of Test = 1 – Probability of committing Type II Error

Specificity of Test = 1 – Probability of committing Type I Error.

Solution 31: Correct Answer is **Option A.** [2]

The 95% confidence interval for the original sample
 $= \bar{x} \pm 1.96 * \sigma / \sqrt{n}$
 $= (56 + 32 + 49 + 57 + 44) / 5 \pm 1.96 * \sqrt{20/5}$
 $= 47.6 \pm 3.92$
 $= (43.68, 51.52).$

Solution 32: Correct Answer is **Option C.** [2]

The 95% prediction interval for the original sample
 $= \bar{x} \pm 1.96 * \sigma * \sqrt{1 + 1/n}$
 $= 47.6 \pm 1.96 * \sqrt{20} * \sqrt{1 + 1/5}$
 $= 47.6 \pm 9.602$
 $= (38.00, 57.20)$

Solution 33: Correct Answer is **Option D.** [2]

The width of the 95% confidence interval
 $= 2 * 1.96 * \sigma / \sqrt{n}$

Higher the value of n, lower will be the width of the 95% confidence interval

The width of the 95% prediction interval
 $= 2 * 1.96 * \sigma * \sqrt{1 + 1/n}$

Higher the value of n, lower will be the width of the 95% prediction interval.

Solution 34: Correct Answer is **Option D.** [2]

Average reduction in BAC levels
 $= (48 + 28 + 42 + 51 + 39) / 5$
 $= 41.60$ basis points

Variance of reduction in BAC levels
 $(\sum (x - \bar{x})^2) / (n-1)$
 $= (40.96 + 184.96 + 0.16 + 88.36 + 6.76) / 4$
 $= 321.2 / 4$
 $= 80.30$

Standard deviation of reduction in BAC levels
 $= \sqrt{80.30}$
 $= 8.96$ basis points.

Solution 35: Correct Answer is **Option D.** [2]

90% confidence interval for the reduction in BAC levels
 $= 41.60 \pm t_4 \text{ (at 5\%)} * 8.96 / \sqrt{5}$
 $= 41.60 \pm 2.132 * 8.96 / \sqrt{5}$
 $= (33.06, 50.14).$

Solution 36: Correct Answer is **Option B.** [2]

Ruling Alliance and Opposition Alliance are equally strong contenders and have an equal chance of winning the poll.

If the Ruling Alliance is re-elected to power, λ is expected to be equal to 100 and if the Opposite Alliance is elected, λ is expected to be equal to 300.

Hence the prior distribution is: $P(\lambda = 100) = 0.5$, $P(\lambda = 300) = 0.5$

Solution 37: Correct Answer is **Option C.** [2]

From Actuarial Tables, it can be seen that the distribution function of the two parameter Pareto distribution is –

$$F(x) = 1 - \left(\frac{\lambda}{(\lambda + x)} \right)^\alpha$$

We know that in the extant case, $\alpha = 1$. Hence,

$$F(x|\lambda) = 1 - \frac{\lambda}{(\lambda + x)}$$

Solution 38: Correct Answer is **Option B.** [2]

Posterior probability that the Ruling Alliance wins the election

$$\begin{aligned} &= f(\lambda=100 | x > 500) \\ &= \frac{(1-F(500|\lambda=100))*P(\lambda=100)}{(1-F(500|\lambda=100))*P(\lambda=100)+(1-F(500|\lambda=300))*P(\lambda=300)} \\ &= \frac{0.1667}{0.1667+0.375} \\ &= 0.307692. \end{aligned}$$

Posterior probability that the Opposition Alliance wins the election

$$\begin{aligned} &= f(\lambda=300 | x > 500) \\ &= \frac{(1-F(500|\lambda=300))*P(\lambda=300)}{(1-F(500|\lambda=100))*P(\lambda=100)+(1-F(500|\lambda=300))*P(\lambda=300)} \\ &= \frac{0.375}{0.1667+0.375} \\ &= 0.692308. \end{aligned}$$

If the stock market index is expected to grow by at least 500 points by the end of the year, Opposition Alliance has a better chance of winning the elections as compared to the Ruling Alliance.

Solution 39: Correct Answer is **Option D.** [2]

We want to find the minimum level of growth in the stock market index which will make the change in government virtually certain.

So, we want to find “a” such that:

$$f(\lambda=300 | x > a) = 1$$

Let us check each of the options.

- A. At $a = 200$, $f(\lambda=300 | x > 200) = 0.64287 < 1$
 B. At $a = 300$, $f(\lambda=300 | x > 300) = 0.66667 < 1$
 C. At $a = 500$, we know that $f(\lambda=300 | x > 500) = 0.692308 < 1$
 D. None of the above – this is the right answer.

Solution 40: Correct Answer is **Option A.** [2]

In case of All or Nothing Loss of 0/1 loss, the loss function is given by the following expression:

$$L(g(\underline{x}), \theta) = 0 \text{ if } g(\underline{x}) = \theta, L(g(\underline{x}), \theta) = 1 \text{ if } g(\underline{x}) \neq \theta$$

Solution 41: Correct Answer is **Option B.** [2]

From Actuarial Tables, we know that for a beta distribution,

$$\mu = \alpha / (\alpha + \beta)$$

$$\begin{aligned} \sigma^2 &= \alpha\beta / ((\alpha + \beta)^2 * (\alpha + \beta + 1)) \\ &= \mu\beta / ((\alpha + \beta) * (\alpha + \beta + 1)) \\ &= \mu\beta / ((\alpha + \beta)^2 + (\alpha + \beta)). \end{aligned}$$

Solution 42: Correct Answer is **Option C.** [2]

Let's check each option and work the value of μ first.

- A. $\mu = 3 / (3+3) = 50\%$ - this is incorrect
 B. $\mu = 9 / (9+1) = 90\%$ - this is incorrect
 C. $\mu = 1 / (1+9) = 10\%$ - this is correct
 D. $\mu = 2 / (2+4) = 66.67\%$ - this is incorrect.

Now, for Option C, let us find out the value of σ^2

$$\begin{aligned} \sigma^2 &= 9 / ((10)^2 * (11)) \\ &= 9 / 1100 - \text{this is correct} \end{aligned}$$

So, the right answer is Option C.

Solution 43: Correct Answer is **Option D.** [2]

$$\begin{aligned} f_{\text{prior}}(\theta) &\propto \theta^{\alpha-1} * (1-\theta)^{\beta-1} \quad 0 < \theta < 1 \end{aligned}$$

Let X be the number of wrong answers given by the AI Chatbot.

$$X | \theta \sim \text{Binomial}(n, \theta)$$

We have observed w wrong answers. So the likelihood function is:

$$\begin{aligned} L(\theta) &= P(X = w | \theta) \\ &= {}^n C_w * \theta^w * (1 - \theta)^{n-w} \\ &\propto \theta^w * (1 - \theta)^{n-w} \end{aligned}$$

Combining the prior distribution and the sample data, we see that:

$$\begin{aligned} f_{\text{posterior}}(\theta) & \propto \theta^{\alpha-1} * (1-\theta)^{\beta-1} * \theta^w * (1-\theta)^{n-w} \\ & \propto \theta^{\alpha+w-1} * (1-\theta)^{\beta+n-w-1} \end{aligned}$$

So, posterior distribution of θ is Beta ($\alpha+w$, $\beta+n-w$).

Solution 44: Correct Answer is **Option A.** [2]

$$\begin{aligned} L(\theta) & = P(x = w | \theta) \\ & = {}^n C_w * \theta^w * (1-\theta)^{n-w} \end{aligned}$$

$$\begin{aligned} \text{Log}L(\theta) & = w * \log\theta + (n-w) * \log(1-\theta) + \text{constant} \end{aligned}$$

$$d/d\theta (\text{Log}L(\theta)) = w/\theta - (n-w)/(1-\theta)$$

Equating this to 0.

$$\begin{aligned} w * (1-\theta) - (n-w) * \theta & = 0 \\ w - w\theta - n\theta + w\theta & = 0 \\ w - n\theta & = 0 \\ \hat{\theta}_{\text{MLE}} & = w / n. \end{aligned}$$

We can also check that $d^2/d\theta^2 (\text{Log}L(\theta)) < 0$ indicating that this is maxima.

Solution 45: Correct Answer is **Option B.** [2]

Mean of the posterior distribution can be expressed in the form of a credibility estimate with the credibility factor $Z = n / (\alpha + \beta + n)$.

$$\begin{aligned} \text{Mean of the posterior distribution} & = (\alpha+w) / (\alpha+w + \beta+n-w) \\ & = (\alpha+w) / (\alpha+\beta+n) \\ & = w / (\alpha + \beta + n) + \alpha / (\alpha + \beta + n) \\ & = n / (\alpha + \beta + n) * (w / n) + (\alpha + \beta) / (\alpha + \beta + n) * \alpha / (\alpha + \beta) \\ & = Z * \hat{\theta}_{\text{MLE}} + (1 - Z) * \mu. \end{aligned}$$

Solution 46: Correct Answer is **Option A.** [2]

For a Poisson distribution, the canonical link function is:

$$g(\mu) = \log \mu$$

Solution 47: Correct Answer is **Option C.** [2]

As mentioned in the question, education score of 1 and mode of earning score of 0 are considered to be base values. Hence, the true value of their coefficients is implicitly covered in the intercept coefficient β_0 .

Solution 48: Correct Answer is **Option C.** [2]

It is given that:

- Mr. Fast is a qualified Chartered Accountant whereas Ms. Steady is an Architect. So, both are professionals.
- Mr. Fast has his own CA practice whereas Ms. Steady works as a freelancer. So, both are self-employed.

Education Score = 1

Mode of Earning Score = 0

$$\begin{aligned} g(\mu) &= \beta_0 + \beta_1 + \beta_2 \\ &= 0.7732 + 0 + 0 \\ &= 0.7732 \end{aligned}$$

Using the given link function, $g(\mu) = \mu$,

$$\begin{aligned} \mu &= 0.7732 \\ &= 77.32\% \end{aligned}$$

Solution 49: Correct Answer is **Option D.**

[2]

The log-likelihood of the Poisson model can be derived as follows:

$$f(y) = \frac{\mu^y e^{-\mu}}{y!}$$

$$l = \sum_{i=1}^n [y_i \log \mu_i - \mu_i - \log y_i!]$$

$$l_f = \sum_{i=1}^n [y_i \log y_i - y_i - \log y_i!]$$

Solution 50: Correct Answer is **Option B.**

[2]

Scaled Deviance

= D

$$= 2(\log L_S - \log L_M)$$

$$= 2(\sum_1^n y_i \log y_i - \sum_1^n y_i - \sum_1^n \log y_i! - \sum_1^n y_i \log \hat{y}_i + \sum_1^n \hat{y}_i + \sum_1^n \log y_i!)$$

$$= 2(\sum_1^n y_i \log(y_i / \hat{y}_i) - \sum_1^n (y_i - \hat{y}_i))$$

$$= 2 * (28.6132 - 16.2375)$$

$$= 24.7514.$$
