Institute of Actuaries of India

CT6 – Statistical Methods

Indicative Solution

November 2008

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 examiner have given credit for any alternative approach or interpretation which they consider to be reasonable

1.

a. Prior distribution: In Bayesian statistics, a population parameter whose value is of interest but is unknown, is modeled by considering its value to be a random variable independent from any sample data from the population. The assumed distribution for this random variable is the 'prior distribution'.

Posterior distribution: When observations have been made on a sample from the population, a better picture of the value of unknown parameter can be found by considering its conditional distribution, given that the particular values in the sample were observed. This is the posterior distribution.

b.

i. The prior density of θ is

$$f(\theta) = \frac{21^{12}\theta^{11}}{\Gamma(12)} e^{-21\theta}, \theta > 0$$

The likelihood function is

$$L(y \mid \theta) = \prod_{i=1}^{12} \theta y_i \exp\left(-\frac{\theta y_i^2}{2}\right)$$

$$\Rightarrow L(y \mid \theta) = \theta^{12}\left(\prod_{i=1}^{12} y_i\right)\exp\left(-\frac{\theta}{2}\sum_{i=1}^{12} y_i^2\right)$$

The posterior density of θ is proportional to the product of the prior density and the likelihood function.

So,

$$f(\theta \mid y) \propto f(\theta).L(y \mid \theta)$$

$$\Rightarrow f(\theta \mid y) \propto \left(\frac{21^{12}\theta^{11}}{\Gamma(12)} e^{-21\theta}\right)\theta^{12}\left(\prod_{i=1}^{12} y_i\right)\exp\left(-\frac{\theta}{2}\sum_{i=1}^{12} y_i^2\right)$$

$$\Rightarrow f(\theta \mid y) = cons * \theta^{23} \exp\left[-\left(21 + \frac{1}{2}\sum y_i^2\right)\theta\right]$$

$$\Rightarrow f(\theta \mid y) = cons * \theta^{23} \exp[-(21 + 42.2086)\theta]$$

$$\Rightarrow f(\theta \mid y) = cons * \theta^{23} \exp[-63.2086\theta]$$

Hence $\theta|\underline{y}$ follows Gamma (24, 63.21).

ii. The Bayesian estimate of θ using the squared error loss function is the mean of the posterior distribution i.e. 24/63.21 = 0.3797.

(Total 8 Marks)

2.

a.

i. Pure strategy: When a player chooses a definite strategy with no chance of picking any other strategy, it is called a pure strategy. This is in contrast with a randomized strategy where a player makes a random selection from a set of strategies, with pre-determined probabilities of selection.

ii. Minimax criterion: The minimax criterion is one of the criteria that may be used by a player to select a strategy. It amounts to minimizing the maximum possible loss, where the minimum is taken over the player's own choices and maximum is taken over different choices of the other player.

In the case of case of randomized strategies, the minimax criterion amounts to minimizing the maximum risk (*expected* loss).

b. There are no dominant strategies and no saddle points, so each player should adopt a randomized strategy.

<u>First consider player A</u>
Suppose that player A adopts strategy I with probability p and strategy II with probability $1 - p$. If player B chooses strategy 1 then player A's expected loss will be $4p - (1 - p) = 5p - 1$. If player B chooses strategy 2 then player A's expected loss will be $-3p + 2(1 - p) = 2 - 5p$. These expressions are equal when $p = 0.3$. Hence player A should choose strategy I with probability 0.3 and strategy II with probability 0.7.

<u>Now consider player B in a similar way</u>
Suppose that player B adopts strategy 1 with probability q and strategy 2 with probability $1 - q$. If player A chooses strategy I then player B's expected gain will be $4q - 3(1 - q) = 7q - 3$. If player A chooses strategy 2 then player B's expected gain will be $-q + 2(1 - q) = 2 - 3q$. These expressions are equal when $q = 0.5$. Hence player B should choose strategy 1 with probability 0.5 and strategy 2 with probability 0.5.

(Even though there is no saddle point, there are minimax pure strategies for the two players. Strategy II is minimax for A and strategy 1 is minimax for B. The implication of there not being a saddle point is that there is always a player who stands to gain by switching the pure strategy if s/he has the knowledge of the opponent's pure strategy. This can go on and on; there is no equilibrium.

Those who derive the minimax pure strategies for the two players and give the full explanation are also given credit.)

c. The loss matrix is given by:

		Statistician (estimated)	
State of nature (actual)		$\hat{\Theta} = 3$	$\hat{\Theta} = 5$
	$\Theta = 3$	0	50
	$\Theta = 5$	20	0

If $\theta = 5$, the probability that the statistician guesses correctly is

$$P(X > k) = \int_k^5 \frac{2x}{25}dx = \frac{25-k^2}{25}.$$

If $\theta = 3$, the probability that the statistician guesses correctly is

$$P(X > k) = \int_0^k \frac{2x}{9}dx = \frac{k^2}{9}.$$

The risk function will be calculated from the size of the loss multiplied by the probability that the statistician guesses wrong. So for the decision rule adopted by the statistician:

$$R(d, \theta = 5) = 20 \times \frac{k^2}{25},$$

$$R(d, \theta = 3) = 50 \times \left(1 - \frac{k^2}{9}\right).$$

The maximum risk will be minimized when these two expressions have the same value i.e.

$$20 \times \frac{k^2}{25} = 50 \times \left(1 - \frac{k^2}{9}\right).$$

$$\Rightarrow k = 2.805.$$

(Total 13 Marks)

3.

a. The advantages of NCD system are:

i. It reduces the policyholder's propensity to claim thereby reducing the number of claims;

ii. It discourages small claims thereby reducing administration costs associated with servicing of such claims;

iii. By offering discount, the insurer is able to retain lower risk customers, who would have otherwise switched to competing insurers.

b.

i. <u>Proportion of Good drivers at each discount level</u>

Let π_0, π_x and π_{2x} denote the proportion of drivers at the 0%, X% and 2X% discount levels respectively.

The steady state vector equation is then

$$[\pi_0, \pi_x, \pi_{2x}] \begin{bmatrix} 0.1 & 0.9 & 0 \\ 0.1 & 0 & 0.9 \\ 0 & 0.1 & 0.9 \end{bmatrix} = [\pi_0, \pi_x, \pi_{2x}]$$

$$\begin{aligned} \Rightarrow 0.1\pi_0 + 0.1\pi_x &= \pi_0 \quad \ldots(1) \\ 0.9\pi_0 + 0.1\pi_{2x} &= \pi_x \quad \ldots(2) \\ 0.9\pi_x + 0.9\pi_{2x} &= \pi_{2x} \quad \ldots(3) \\ \pi_0 + \pi_x + \pi_{2x} &= 1 \quad \ldots(4) \end{aligned}$$

From equation (3), $\pi_x = \dfrac{\pi_{2x}}{9}$

From equation (1) $0.1\pi_x = 0.9\pi_0 \quad \Rightarrow \pi_0 = \dfrac{\pi_{2x}}{81}$

From equation (4),

$$\frac{\pi_{2x}}{81} + \frac{\pi_{2x}}{9} + \pi_{2x} = 1$$

$$\Rightarrow \pi_{2x} = \frac{81}{91}, \pi_x = \frac{9}{91}, \pi_0 = \frac{1}{91}$$

Since there are 7735 good drivers altogether, there must be 85, 765 and 6885 good drivers on the average at the 0%, *X*% and 2*X*% discount levels respectively.

Proportion of Bad drivers at each discount level

Repeating the above calculations for bad drivers, the steady state vector equation is

$$[\pi_0, \pi_x, \pi_{2x}] \begin{bmatrix} 0.2 & 0.8 & 0 \\ 0.2 & 0 & 0.8 \\ 0 & 0.2 & 0.8 \end{bmatrix} = [\pi_0, \pi_x, \pi_{2x}]$$

$$\begin{aligned} \Rightarrow 0.2\pi_0 + 0.2\pi_x &= \pi_0 \quad \ldots(1) \\ 0.8\pi_0 + 0.2\pi_{2x} &= \pi_x \quad \ldots(2) \\ 0.8\pi_x + 0.8\pi_{2x} &= \pi_{2x} \quad \ldots(3) \\ \pi_0 + \pi_x + \pi_{2x} &= 1 \quad \ldots(4) \end{aligned}$$

From equation (3), $\pi_x = \dfrac{\pi_{2x}}{4}$.

From equation (1) $0.2\pi_x = 0.8\pi_0 \quad \Rightarrow \pi_0 = \dfrac{\pi_{2x}}{16}$.

From equation (4),

$$\frac{\pi_{2x}}{16}+\frac{\pi_{2x}}{4}+\pi_{2x}=1.$$

$$\Rightarrow \pi_{2x}=\frac{16}{21}, \pi_x=\frac{4}{21}, \pi_0=\frac{1}{21}.$$

Since there are 2205 bad drivers altogether, there must be 105, 420 and 1680 bad drivers on the average at the 0%, X% and $2X$% discount levels respectively.

ii. The average premium paid by the good drivers is:

$$\frac{P}{91}+\frac{9}{91}\left(1-\frac{X}{100}\right)P+\frac{81}{91}\left(1-\frac{2X}{100}\right)P=\left(1-\frac{171X}{9100}\right)P.$$

The average premium paid by the bad drivers is:

$$\frac{P}{21}+\frac{4}{21}\left(1-\frac{X}{100}\right)P+\frac{16}{21}\left(1-\frac{2X}{100}\right)P=\left(1-\frac{36X}{2100}\right)P.$$

iii. If the average premium paid by good drivers is equal to half the average premium paid by bad drivers then

$$\left(1-\frac{171X}{9100}\right)P=\frac{1}{2}\left(1-\frac{36X}{2100}\right)P \quad \Rightarrow X=48.9\%.$$

Comment: The computed discount level to produce reasonable contrast in the premium charged to good and bad drivers is too large to implement in practice. This is one of the limitations of the NCD system.

(Total 13 Marks)

4.

i. $V(S) = E(N)V(X) + V(N)[E(X)]^2$

ii. N has Geometric distribution. So,
$E(N) = q/p$,
$V(N) = q/p^2$

X has Exp(λ) distribution. So,
$E(X) = 1/\lambda$,
$V(X) = 1/\lambda^2$

$V(S) = E(N)V(X) + V(N)[E(X)]^2$

$$= \frac{q}{p} \times \frac{1}{\lambda^2} + \frac{q}{p^2} \times \frac{1}{\lambda^2}$$

$$= \frac{q(p+1)}{\lambda^2 p^2}.$$

iii. If c is the annual rate of premium, then the premium collected over time period of length t is ct.
Expected aggregate claim during the same period = $E(N)E(X) = \theta t/\lambda$.
Equating the two, we have $c = \theta/\lambda$.
Surplus at time t, $U(t) = U + ct - S(t)$, where
$S(t) = X_1 + X_2 + \cdots + X_{N(t)}$;
$X_1 + X_2 + \cdots + X_n$ is gamma(λ,n), $N(t)$ is Poisson (θt).
Therefore, probability of ruin

$$= P(U + ct - S(t) < 0)$$

$$= P(S(t) > U + ct)$$

$$= \sum_{n=0}^{\infty} P(S(t) > U + ct \mid N(t) = n) P(N(t) = n)$$

$$= \sum_{n=1}^{\infty} P(N(t) = n) P\left(S(t) > U + \frac{\theta t}{\lambda} \mid N(t) = n \right)$$

$$= \sum_{n=1}^{\infty} \left(\frac{(\theta t)^n e^{-\theta t}}{n!} \right) \int_{U+\theta t/\lambda}^{\infty} \frac{x^{n-1} \lambda^n e^{-\lambda x}}{(n-1)!} dx.$$

(Total 8 Marks)

5.

a.

i. A covariate is a variable in a generalized linear model about which one has some information collected and whose values may carry some information about the main variable of interest – the response variable. One often tries to predict values of the response variable based on the values taken by the covariates. Covariates may take numerical or categorical values.

ii. In the linear model, the mean response variable is expressed directly as a linear function involving covariates. In a generalized linear model, one expresses some function of the mean response in the linear form. This function is called the link function. In the special case of the linear model, the link function is the identity function.
The linear form is linear in the regression parameters associated with the covariates. When the distribution of the response belongs to the exponential family, there is a natural form of the link function called the canonical link function, that simplifies computation.

b. A probability distribution is said to belong to an exponential family if its probability density function is of the form

$$f(y;\theta,\varphi) = \exp[\frac{y\theta - b(\theta)}{a(\varphi)} + c(y,\varphi)],$$

Where a, b and c are functions, θ is the natural parameter and φ is another parameter (which often turns out to be the scale parameter).

c. Solution

i.

$f(x) = \lambda e^{-\lambda x}$

$= \exp[\log(\lambda) - \lambda x]$

$= \exp[x(-\lambda) - (-\log\lambda)]$

This conforms to the exponential family with $\theta = -\lambda$, $b(\theta) = -\log\lambda$, $a(\varphi) = 1$ and $c(y,\varphi) = 0$.

(The minus signs in the expressions of θ and $b(\theta)$ can be absorbed in $a(\varphi)$ also.)

ii.

$$f(x) = crx^{r-1}\exp(-cx^r)$$

$$\Rightarrow f(x) = \exp(\log(crx^{r-1}).\exp(-cx^r)$$

$$\Rightarrow f(x) = \exp[\log(c) + \log(r) + (r-1)\log(x) - cx^r].$$

Here the x term appears in power not equal to unity. There is no way we can write it as a first degree function in x because $r \neq 1$. Hence the distribution does not belong to the exponential family.

iii.

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

$= \exp[\log(1 - q)q^x)$

$= \exp[\log(1 - q) + x\log(q)]$

$= \exp[x\log(q) - (-\log(1 - q)]$

Which conforms to exponential family with $\theta = \log(q)$, $b(\theta) = -\log(1 - q)$, $a(\varphi) = 1$ and $c\ (x, \varphi) = 0$.

(Total 10 Marks)

6.

a.

i. Recast the incremental claims paid data in 'development of claims' pattern

(Amounts in '000s)

		Development year				
		0	1	2	3	4
Accident year	1997	28,791	22,063	2,805	378	78
	1998	27,620	2,310	17,725	8,256	
	1999	26,935	11,925	9,872		
	2000	36,661	9,222			
	2001	18,619				

Cumulative claims paid amounts

		Development year				
		0	1	2	3	4
Accident year	1997	28,791	50,854	53,659	54,037	54,115
	1998	27,620	29,930	47,655	55,911	
	1999	26,935	38,860	48,732		
	2000	36,661	45,883			
	2001	18,619				
	Dev fct	1.37931	1.25410	1.08522	1.00144	1.0000
	Cum dev fct	1.87992	1.36294	1.08679	1.00144	1.0000

O/S claims reserve = 54115*(1-1) + 55911*(1.00144-1) + 48732*(1.08679-1) + 45883*(1.36294-1) + 18619*(1.87992-1)
= 37.346 million

ii. For inflation, there is an implied assumption that 'a weighted average of past inflation will be repeated in the future'

b. <u>Inflation adjusted Chain Ladder Method</u>

All payments will need to be converted to mid 2001 values. For this the following index table is used
Incremental claims paid at mid 2001 values

Year	Inflation index
1997-2001	1.05*1.07*1.06*1.09 =1.2981
1998-2001	1.07*1.06*1.09 =1.2363
1999-2001	1.06*1.09 = 1.1554
2000-2001	1.09

Applying Chain Ladder method to cumulative inflated amounts

	Development year					
		0	1	2	3	4
Accident year	1997	37,373	64,649	67,890	68,302	68,380
	1998	34,146	36,815	56,135	64,391	
	1999	31,121	44,119	53,991		
	2000	39,960	49,182			
	2001	18,619				
	Dev fct	1.3658	1.2228	1.0699	1.0011	1.000
	Cum dev fct	1.7888	1.3097	1.0711	1.0011	1.000

Projected cumulative paid amounts

	Development year					
		0	1	2	3	4
Accident year	1997					
	1998					64,465
	1999				57,765	57,831
	2000			60,140	64,344	64,417
	2001		25,430	31,096	33,269	33,307

Projected year by year amounts at mid 2001 values

	Development year					
		0	1	2	3	4
Accident year	1997					
	1998					74
	1999				3,774	66
	2000			10,958	4,204	73
	2001		6,811	5,666	2,173	38

Adjust for future claims inflation and discount for interest earned. Assume payments are made mid-year

Due in year	Amount	Adj. for inflation	Discount factor	Present value @ 31/12/2001
2002	21,617	1.06	$1.08^{-1/2}$	22,047
2003	9,936	1.06*1.06	$1.08^{-3/2}$	9,947
2004	2,246	1.06*1.06*1.05	$1.08^{-5/2}$	2,186
2005	38	1.06*1.06*1.05*1.05	$1.08^{-7/2}$	36

Summing the last column, the discounted O/S claims reserve is Rs 34.218 million

(Total 16 Marks)

7.

a. Let S be the aggregate claim amount before reinsurance.
Then $S = X_1+X_2+\dots\dots\dots+X_N$, where X has Pareto(α, λ) with $\alpha = 3$ and $\lambda = 1000$

So, $E(X) = \lambda/(\alpha - 1) = 500$
$V(X) = \alpha\lambda^2/(\alpha - 2)(\alpha - 1)^2 = 750{,}000$

Since the Poisson parameter is μ, we have,
$E(S) = 500\mu$ and $V(S) = 1000000\mu$

The insurer's expected profit without reinsurance is equal to the premiums minus expected claims. But if a loading factor of 0.2 is in use then the total premium is $=1.2*500\mu = 600\mu$ and the expected profit is $600\mu - 500\mu = 100\mu$

b. Now consider the effect of reinsurance. Let Y denote the amt paid by the insurer

$$Y = X \quad \text{if } X < 1000,$$
$$=1000 \quad \text{if } X \geq 1000.$$

Let Z denote the amount paid by the reinsurer.

$$Z = 0 \quad \text{if } X < 1000,$$
$$= X - 1000 \quad \text{if } X \geq 1000.$$

So for the reinsurer, total aggregate claims are given by $S_R = Z_1 + Z_2 + \cdots + Z_N$.

The reinsurance premium is given by $1.3E(S_R)$, where $E(S_R) = E(N)E(Z) = \mu\, E(Z)$.

$$E(Z) = \int_{1000}^{\infty} (x-1000) * \frac{3*1000^3}{(1000+x)^4} dx\ .$$

Putting $u = x - 1000$ in the integral,

$$E(Z) = \int_0^{\infty} u * \frac{3*1000^3}{(2000+u)^4} du$$

$$= \left(\frac{1000}{2000}\right)^3 \int_0^{\infty} u * \frac{3*2000^3}{(2000+u)^4} du.$$

The integral being the mean of Pareto(3,2000) distribution, $E(Z) = 125$

$E(S_R) = 1.3*125\mu = 162.5\mu$

Since the expected recovery from the reinsurer is $E(S - S_R)$, and the expected profit with the reinsurance is $600\mu - 162.5\mu - E(S - S_R)$. As

$E(S - S_R) = E(S) - E(S_R) = 375\mu$.

So the expected profit is 62.5μ and the percentage reduction in the expected profit is 37.5%.

(Total 8 Marks)

8.

a. For $-3 < x < 3$,

$$f(x) = \varphi(x \mid -3 < x < 3) = \frac{\varphi(x)}{P(-3 < X < 3)}$$

$$= \frac{\varphi(x)}{\phi(3) - \phi(-3)} = \frac{\varphi(x)}{2\phi(3) - 1}$$

$$\Rightarrow k = (2(.9987) - 1)^{-1} = .9973^{-1} = 1.0027.$$

b. Let the linear transformation needed to $Y = cX + d$. Y~Uniform(–3,3).
We may have $Y = -3$ for $X = 0$ and $Y = 3$ for $X = 1$.
So, $d = -3$ and $c + d = 3$, i.e. $c = 6$.
So the transformation needed is $6X - 3$.
(The answer $Y = 3 - 6X$ fetches full credit.)

c. Step-wise algorithm

i. Generate a random variate x from Uniform (0,1).
ii. Transform x using $y = 6x - 3$ to get a random variate from Uniform (– 3,3).
iii. Generate another random variate z from Uniform (0,1).
iv. If $z \le \sqrt{2\pi} f(x)$ then accept y as the required random variate, otherwise reject y and go to step (i).

OR

i. Generate a random variate x from Uniform (0,1).
ii. Transform x using $y = 6x - 3$ to get a random variate from Uniform (– 3,3).
iii. Generate another random variate z from Uniform $(0,1/\sqrt{2\pi})$.
iv. If $z \le f(x)$ then accept y as the required random variate, otherwise reject y and go to step (i).

OR

i. Generate random variates from the untruncated normal distribution using the usual Box-Muller transformation
$z_1 = \sqrt{-2\log(u_1)}\cos(2\pi u_2); \quad z_2 = \sqrt{-2\log(u_1)}\sin(2\pi u_2),$
OR the polar transformation
$$z_1 = \sqrt{\frac{-2\log((u_1^2+u_2^2))}{u_1^2+u_2^2}}(2u_1-1); \quad z_2 = \sqrt{\frac{-2\log((u_1^2+u_2^2))}{u_1^2+u_2^2}}(2u_2-1),$$
Provided that $u_1^2+u_2^2 \le 1$ (try again if $u_1^2+u_2^2 > 1$).

ii. Reject and repeat if the absolute value of the generated number is more than 3.

(Total 8 Marks)

9.

a. The partial autocorrelation function ϕ_k is defined as the conditional correlation of X_t with X_{t-k} given $X_{t-1},\ldots, X_{t-k+1}$.
It may be derived as the coefficient ϕ_k in the problem to:
Min $E[(X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \cdots - \phi_k X_{t-k})^2]$.

b. For k>2:
$E[(X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \cdots - \phi_k X_{t-k})^2]$
$= E[(\alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \varepsilon_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \cdots - \phi_k X_{t-k})^2]$

$= E[E\{(\alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \varepsilon_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} -$
$\cdots - \phi_k X_{t-k})^2 | X_{t-1},\ldots, X_{t-k+1}\}]$
$= E[E\{\varepsilon_t^2 | X_{t-1},\ldots, X_{t-k+1}\}]$
$+2E[E\{\varepsilon_t (\alpha_1 X_{t-1} + \alpha_2 X_{t-2} - \phi_1 X_{t-1} - \phi_2 X_{t-2} -$
$\cdots - \phi_k X_{t-k}) | X_{t-1},\ldots, X_{t-k+1}\}]$
$+E[E\{(\alpha_1 X_{t-1} + \alpha_2 X_{t-2} - \phi_1 X_{t-1} - \phi_2 X_{t-2} -$
$\cdots - \phi_k X_{t-k})^2 | X_{t-1},\ldots, X_{t-k+1}\}]$

$= E(\varepsilon_t^2) + E[E\{(\alpha_1 X_{t-1} + \alpha_2 X_{t-2} - \phi_1 X_{t-1} - \phi_2 X_{t-2} -$
$\cdots - \phi_k X_{t-k})^2 | X_{t-1},\ldots, X_{t-k+1}\}]$
The second term is minimized when $\phi_1=\alpha_1$, $\phi_2=\alpha_2$, $\phi_3=\ldots=\phi_k=0$.
Therefore, $\phi_k=0$.

(Total 6 Marks)

10.

a. The characteristic equation is
$(1 - 11t/6 + t^2 - t^3/6) = 0$, i.e., $t^3 - 6t^2 + 11t - 6 = 0$.
As 1 is a root of the characteristic equation, $(t-1)$ is one of the factors of the LHS.
We will need to get the factor $(t-1)$ from this equation.
$t^2(t-1) + t^2 - 6t^2 + 11t - 6 = 0.$
<=> $t^2(t-1) - 5t^2 + 11t - 6 = 0.$
<=> $t^2(t-1) - 5t(t-1) - 5t + 11t - 6 = 0.$

$\Leftrightarrow$ $t^2(t-1) - 5t(t-1) + 6t - 6 = 0.$
$\Leftrightarrow$ $t^2(t-1) - 5t(t-1) + 6(t-1) = 0.$
$\Leftrightarrow$ $(t-1)(t^2 - 5t + 6) = 0.$
$\Leftrightarrow$ $(t-1)(t-2)(t-3)=0.$
$\Leftrightarrow$ $t = 1, 2, 3.$

b. Although two of the roots are greater than 1, the process is not stationary because the third root is equal to 1.

c. $p = 2, d = 1, q = 0.$

d. If we define $Y_t = X_t - X_{t-1}$, then the stationary process Y_t is AR(2), and it satisfies the equation $Y_t = 1 + \frac{5}{6}Y_{t-1} - \frac{1}{6}Y_{t-2} + \varepsilon_t$.

Taking expectation of both sides of the equation, the mean μ of the process satisfies the equation

$\mu = 1 + \frac{5}{6}\mu - \frac{1}{6}\mu$.

Therefore, $E(Y_t) = \mu = 3$

e. The mean of X_t does not exist.
(Full credit for saying that the mean is infinity or that it depends on one initial value.)

f. The forecast of ε_t is 0. Given that $Y_{99} = X_{99} - X_{98} = -0.8$ and $Y_{100} = X_{100} - X_{99} = -0.4$, the forecast of Y_{101} is $1 + \frac{5}{6}Y_{100} - \frac{1}{6}Y_{99} = 0.8$.
Therefore, the forecast of X_{101} is $X_{100} + 0.8 = 2$.

(Total 10 Marks)

([Total 100 Marks)

*************END*************