Stat 24400 Homework 3 Solution

Jan 27, 2016

Total points: 100

1. **[10pts] Rice 4.78**

   Suppose $f(x)$ is a probability density function and symmetric about zero, then it is an even function ($f(x) = f(-x)$). Let k be an odd number then $xkf(x)$ is an odd function and hence

   $$E(X^k) = \int_{-\infty}^{\infty} x^k f(x)dx = 0$$

   So all odd moments of X is 0. As a result,

   $$skewness = E((X - E(X))^3 = E(X^3) = 0$$

   Grading Scheme: 5 pts for proving $E(X) = 0$, and 5 pts for the rest. Marks are assigned based on the progress made through.

2. **[30 pts]**

   (a)

   $$\int_0^1 \int_0^1 f(x,y)dxdy = \frac{4}{5}\int_0^1 \int_0^1 (x + y + xy)dxdy = \frac{4}{5}\int_0^1 \frac{1}{2} + \frac{3}{2}ydy = 1$$

   (b)

   $$f_Y(y) = \int_0^1 f(x,y)dx = \int_0^1 (x + y + xy)dx = \frac{2}{5}(3y + 1)$$

   for $0 < y < 1$.

   (c)

   $$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{2(x + y + xy)}{3y + 1}$$

   Then

   $$f_{X|Y=0.5}(x|y = 0.5) = \frac{2(x + 0.5 + x/2)}{1.5 + 1} = \frac{2}{5}(3x + 1)$$

   for $0 < x < 1$.

   (d) Note that $f(x,y)$ is symmetric in x and y, hence X and Y have the same distribution (they are not independent though).

   $$E(X) = E(Y) = \frac{2}{5}\int_0^1 (3y + 1)ydy = \frac{3}{5}$$

$$E(X^2) = E(Y^2) = \frac{2}{5}\int_0^1 (3y+1)y^2 dy = \frac{13}{30}$$

$$Var(X) == E(X^2) - (E(X))^2 = \frac{11}{150}$$

$$E(XY) = \frac{4}{5}\int_0^1 \int_0^1 xy(x+y+xy)dxdy = \frac{4}{5}\int_0^1 \frac{1}{3}y + \frac{5}{6}y^2 dy = \frac{16}{45}$$

$$Cov(X,Y) = E(XY) - E(X)E(Y) = -\frac{1}{225}$$

(e)

$$P(0.2 \leq X \leq 0.5, 0.4 \leq Y \leq 0.8) = \frac{4}{5}\int_{0.2}^{0.5}\int_{0.4}^{0.8} (x+y+xy)dydx$$
$$= \frac{4}{5}\int_{0.2}^{0.5} (0.64x + 0.24)dx = 0.11136$$

(f)

$$P(X+Y \leq 1) = \frac{4}{5}\int_0^1 \int_0^{1-x} (x+y+xy)dydx = \frac{2}{5}\int_0^1 (x^3 - 3x^2 + x + 1)dx = \frac{3}{10}$$

Grading Scheme: 5 pts for each sub-problem. For each sub-problem, deduct 2 pts if answer is wrong but steps are valid. For (b) and (c), deduct 2 pts is domain is not given.

3. **[20 pts] Rice 4.81 and 4.82**

Let $X \sim Ber(p)$, then
$$M_X(t) = E(e^{tX}) = e^t p + (1-p)$$

For $k > 1$, the k-th derivatives are identical in this case, i.e.

$$M_X^{(k)}(t) = e^t p$$

As a result,
$$E(X) = M_X^{(1)}(0) = p$$
$$Var(X) = E(X^2) - (E(X))^2 = M_X^{(2)}(0) - (M_X^{(1)}(0))^2 = p(1-p)$$
$$E(X^3) = M_X^{(3)}(0) = p$$

Now let $Y \sim B(n, p)$, then $Y = \sum_{i=1}^n X_i$, where X_i are i.i.d. *Bernoulli*(p) r.v.

$$M_Y(t) = E(e^{tY}) = E(e^{\sum_{i=1}^n tX_i}) = \prod_{i=1}^n E(e^{tX_i}) = M_X^n(t) = (e^t p + 1 - p)^n$$

where the 3rd equality follows from the independence of Xi's and positiveness of the exponential function. You could still evaluate the derivative of $M_Y(t)$ at 0 to calculate the moments. Alternatively,

$$E(Y) = \sum_{i=1}^n E(X_i) = np$$

$$Var(Y) = \sum_{i=1}^n Var(X_i) = np(1-p)$$

Again, the covariance terms vanishes because X_i's are independent.

Grading Scheme: 5 pts Bernoulli MGF, 5 pts for Bernoulli moments, 5 pts for Binomial MGF, 5 pts for Binomial moments.

4. **[20pts]**
Assume $X \sim Gamma(\alpha, \lambda)$, then its mgf is

$$M_X(t) = (1 - \frac{t}{\lambda})^{-\alpha}$$

Since $E(X) = \frac{\alpha}{\lambda}$,$Var(X) = \frac{\alpha}{\lambda^2}$, then the standardized r.v. is

$$Y = \frac{X - \frac{\alpha}{\lambda}}{\sqrt{\frac{\alpha}{\lambda^2}}} = \frac{\lambda}{\sqrt{\alpha}}X - \sqrt{\alpha}$$

Therefore,

$$M_Y(t) = e^{-\sqrt{\alpha}t} M_X(\frac{\lambda}{\sqrt{\alpha}}t) = e^{-\sqrt{\alpha}t}(1 - \frac{t}{\sqrt{\alpha}})^{-\alpha}$$

To see the behavior when $\alpha \to \infty$, we take the log and then use Taylor's expansion up to the 2nd order.

$$log(M_Y(t)) = -\sqrt{\alpha}t - \alpha log(1 - \frac{t}{\sqrt{\alpha}}) = -\sqrt{\alpha}t - \alpha(-\frac{t}{\sqrt{\alpha}} + \frac{t^2}{2\alpha}) + o(\frac{1}{\alpha^3}) = \frac{t^2}{2} + o(1)$$

which is exactly $log(e^{t^2/2}) = log(M_Z(t))$, where $Z \sim N(0, 1)$.

Grading Scheme: 5 pts for $M_X(t)$, 5 pts for $M_Y(t)$, 10 pts for the remaining deduction.

5. **[10 pts]**

The prior of cure rate θ is $Beta(2, 1)$, hence

$$f(\theta) \propto \theta$$

The probability model is

$$X \sim B(3, \theta)$$

So posterior of θ can be calculated as

$$f(\theta|X = k) \propto f(X = k|\theta)f(\theta) \propto \theta^k(1-\theta)^{3-k}\theta = \theta^{k+1}(1-\theta)^{3-k}$$

As a result,

$$\theta|X = k \sim Beta(k + 2, 4 - k)$$

Then

$$P(\theta \leq 0.2|X = k) = P(Beta(k + 1, 4 - k) \leq 0.2)$$

$$E(\theta|X = k) = \frac{k+2}{6}$$

The numerical values for each k are tabulated as follows:

k	$P(\theta \leq 0.2\|X = k)$	$E(\theta\|X = k)$
0	0.26	0.33
1	0.058	0.5
2	0.0067	0.67
3	0.00032	0.83

A simple R function that enables the calculation is

```
Q4=function(k){
pr=pbeta(0.2,k+2,4-k)
ex=(k+2)/6
return(list(pr,ex))
}
```

Grading Scheme: 8 pts for posterior distribution, 2 pts for numerical calculation.

6. [10 pts]
Denote the probability that the sun rises on a given day as θ. The prior of θ is uniform, so

$$f(\theta) \propto 1$$

Denote $X_i, i = 1, \cdots, n$ as the random variable that takes on value 1 if sun rises and 0 otherwise. So $X_i \sim Ber(\theta)$.
The posterior distribution of θ after observing n days of sunrise is therefore

$$f(\theta|X_1 = 1, \cdots, X_n = 1) \propto f(\theta)f(X_1 = 1, \cdots, X_n = 1|\theta) \propto \prod_{i=1}^{n} f(X_i|\theta) = \theta^n$$

So

$$\theta|X_1 = 1, \cdots, X_n = 1 \sim Beta(n + 1, 1)$$

So the posterior expectation of θ is

$$E(\theta|X_1 = 1, \cdots, X_n = 1) = \frac{n + 1}{n + 2}$$

Grading Scheme: 6 pts for posterior distribution, 4 pts for expectation.