PROBABILITY AND STATISTICS I
HOMEWORK VIII

Jianyu Dong 2019511017

April, 13 2021

1

1.1

By the definition, the p.m.f. of misprints on a particular page is

$$p_X(x) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!}, & for\ x = 0, 1, 2, \ldots, \\ 0, & otherwise. \end{cases}$$

So we could get that the probability that a particular page will contain no misprints is

$$p_X(0) = e^{-\lambda}$$

1.2

By the definition, we have that the p.m.f. of misprints on one page is

$$p_X(x) = \begin{cases} \frac{e^{-\lambda}(\lambda)^x}{x!}, & for\ x = 0, 1, 2, \ldots, \\ 0, & otherwise. \end{cases}$$

So let A ={at least m pages contain more that k misprints} and B={at most (n-m) pages contain less than or equal to k misprints } the we could determine

$$P(A) = P(B) = \sum_{j=0}^{n-m} \binom{n}{j} \left(\sum_{i=0}^{k} p_X(i)\right)^j \left(\sum_{i=k+1}^{\infty} p_X(i)\right)^{n-j} = \sum_{j=0}^{n-m} \binom{n}{j} \left(\sum_{i=0}^{k} \frac{e^{-\lambda}\lambda^i}{i!}\right)^j \left(\sum_{i=k+1}^{\infty} \frac{e^{-\lambda}\lambda^i}{i!}\right)^{n-j}$$

2

Since X_1 and X_2 follow Possion distribution with parameters λ_1 and λ_2, by the definition, we could get tha probability distribution functions of X_1 and X_2 are

$$P_{X_1}(x_1) = \frac{e^{-\lambda_1}\lambda_1^{x_1}}{x_1!}\ for\ x_1 = 0, 1, 2, \ldots,\ P_{X_2}(x_2) = \frac{e^{-\lambda_2}\lambda_2^{x_2}}{x_2!}\ for\ x_2 = 0, 1, 2, \ldots$$

zero elsewhere.

By the definition, we have the conditional probability is

$$P(X_1 = x_1 \mid X_1 + X_2 = k) = \frac{P(X_1 = x_1 \cap X_1 + X_2 = k)}{P(X_1 + X_2 = k)}.$$

Since X_1 and X_2 are independent, we could get that

$$P(X_1 = x_1 \cap X_1 + X_2 = k) = P(X_1 = x_1)P(X_2 = k - x_1) = \frac{e^{-\lambda_1}\lambda_1^{x_1}}{x_1!}\frac{e^{-\lambda_2}\lambda_2^{k-x_1}}{(k - x_1)!}$$

and $X_1 + X_2$ follows Possion distribution with parameter $\lambda_1 + \lambda_2$, so we could get that

$$P(X_1 + X_2 = k) = \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1 + \lambda_2)^k}{k!}.$$

So we get the conditional distribution of X_1 given $X_1 + X_2 = k$, $k \in Z^+$ is

$$P(X_1 = x_1 \mid X_1 + X_2 = k) = \binom{k}{x_1}\frac{\lambda_1^{x_1}\lambda_2^{k-x_1}}{(\lambda_1 + \lambda_2)^k},\ for\ x_1 = 0, 1, 2, \ldots, k.$$

3

Let Y be the total number of items produced by the machine, according to the question, Y follows Possion distribution with parameter λ, so the p.d.f. of Y is

$$p_Y(y) = \frac{e^{-\lambda}\lambda^y}{y!},\ for\ y = 0, 1, 2, \ldots.$$

zero elsewhere.

Let X be the number of defective items produced by the machine, according to the question, we could get the conditional p.d.f. of X is

$$P(X = x \mid Y = y) = \binom{y}{x} p^x (1-p)^{y-x},\ for\ x = 0, 1, 2, \ldots, y.$$

By the theorem, we have that the joint c.d.f. of X and Y is

$$p_{X,Y}(x, y) = p_Y(y) P(X = x \mid Y = y) = \frac{e^{-\lambda}(\lambda p)^x}{x!} \frac{[\lambda(1-p)]^{y-x}}{(y-x)!}.$$

So the marginal distribution of the number of defective items produced by the machine is

$$p_X(x) = \sum_{y=x}^{\infty} \frac{e^{-\lambda}(\lambda p)^x}{x!} \frac{[\lambda(1-p)]^{y-x}}{(y-x)!} = \frac{e^{-\lambda}(\lambda p)^x}{x!} \sum_{t=0}^{\infty} \frac{[\lambda(1-p)]^t}{t!} = \frac{e^{-p\lambda}(p\lambda)^x}{x!},\ for\ x = 0, 1, 2, \ldots$$

zero elsewhere.

4

Since Y have the binomial distribution with parameters n and $p \in (0, 1)$, by the definition, we have the p.m.f. of Y is

$$p_Y(y) = P(Y = y) = \binom{n}{y} p^y (1-p)^y,\ for\ y = 0, 1, 2, \ldots, n.$$

zero elsewhere.

Since X_m have the hypergeometric distribution with parameters $(A_m + B_m, A_m, n)$, by the definition, we have the p.m.f. of X is

$$p_X(x) = P(X = x) = \frac{\binom{A_m}{x}\binom{B_m}{n-x}}{\binom{A_m+B_m}{n}},\ for\ x = 0, 1, 2, \ldots, n.$$

zero elsewhere.

Thus, we could get that

$$\begin{aligned} P(X_m = x) &= \frac{A_m!}{x!(A_m - x)!} \frac{B_m!}{(n-x)!(B_m - n + x)!} \frac{n!(A_m + B_m - n)!}{(A_m + B_m)!} \\ &= \binom{n}{x} \frac{\prod_{i=1}^{x}(A_m - x + i) \prod_{j=1}^{n-x}(B_m - n + x + j)}{\prod_{k=1}^{n}(A_m + B_m + k)} \end{aligned}$$

Since

$$\lim_{m\to\infty} A_m = \infty,\ \lim_{m\to\infty} B_m = \infty,\ \lim_{m\to\infty} \frac{A_m}{A_m + B_m} = p$$

we could get that

$$\begin{aligned} \lim_{m\to\infty} P(X_m = x) &= \lim_{m\to\infty} \binom{n}{x} \frac{(A_m)^x (B_m)^{n-x}}{(A_m + B_m)^n} = \lim_{m\to\infty} \binom{n}{x} \left(\frac{A_m}{A_m + B_m}\right)^x \left(1 - \frac{A_m}{A_m + B_m}\right)^{n-x} \\ &= \binom{n}{x} p^x (1-p)^{n-x} = P(Y = x). \end{aligned}$$

So we get that

$$\lim_{m\to\infty} \frac{P(Y=x)}{P(X_m=x)} = 1.$$

5

Since Y have the Possion distribution with mean λ, by the definition, we have the p.m.f. of Y is

$$p_Y(y) = P(Y=y) = \frac{e^{-\lambda}\lambda^y}{y!},\ for\ y = 0,1,2,\dots$$

zero elsewhere.

Since X_m have the hypergeometric distribution with parameters (A_m+B_m, A_m, n_m), by the definition, we have the p.m.f. of X is

$$p_X(x) = P(X=x) = \frac{\binom{A_m}{x}\binom{B_m}{n_m-x}}{\binom{A_m+B_m}{n_m}},\ for\ x = 0,1,2,\dots,n.$$

zero elsewhere.

Thus, we could get that

$$\begin{aligned} P(X_m=x) &= \frac{A_m!}{x!(A_m-x)!}\frac{B_m!}{(n_m-x)!(B_m-n_m+x)!}\frac{n_m!(A_m+B_m-n_m)!}{(A_m+B_m)!} \\ &= \frac{1}{x!}\frac{A_m!}{(A_m-x)!}\frac{n_m!}{(n_m-x)!}\frac{\prod_{i=1}^{n_m-x}(B_m-n_m+x+i)}{\prod_{j=1}^{n_m}(A_m+B_m+j)} \end{aligned}$$

Since

$$\lim_{m\to\infty} A_m = \infty,\ \ \lim_{m\to\infty} B_m = \infty,\ \ \lim_{m\to\infty} n_m = \infty,\ \ \lim_{m\to\infty} \frac{n_m A_m}{A_m+B_m} = \lambda$$

we could get that

$$\begin{aligned} \lim_{m\to\infty} P(X_m=x) &= \lim_{m\to\infty} \frac{1}{x!}(n_m A_m)^x \frac{(B_m)^{n_m-x}}{(A_m+B_m)^{n_m}} = \lim_{m\to\infty} \frac{1}{x!}\left(\frac{n_m A_m}{A_m+B_m}\right)^x \left(1-\frac{\frac{n_m A_m}{A_m+B_m}}{n_m}\right)^{n_m-x} \\ &= \frac{e^{-\lambda}\lambda^x}{x!} = P(Y=x). \end{aligned}$$

So we get

$$\lim_{m\to\infty} \frac{P(Y=x)}{P(X_m=x)} = 1$$

6

6.1

According to the problem, we have the conditional p.d.f. of X is

$$P(X=x \mid P=p) = (1-p)^x p,\ for\ x = 0,1,2\dots$$

By Bayes' Theorem, we could get that the conditional p.d.f. of P given X=12 is

$$f(p \mid X=12) = \frac{f(p)P(X=12 \mid P=p)}{\int_0^1 f(p)P(X=12 \mid P=p)} = \frac{10(1-p)^{21}p}{\int_0^1 10(1-p)^{21}p\,dp} = 506(1-p)^{21}p,\ for\ 0<p<1,$$

zero elsewhere.

6.2

To find the value p so that the conditional pdf of P achieves the absolute maximum, we only need to calculate that

$$\frac{df(p \mid X = 12)}{dp} = 506 \left[(1-p)^{21} - 21(1-p)^{20}p\right] = 0.$$

So $p = \frac{1}{22}$ then the conditional p.d.f. of P achieves the absolute maximum.

7

Since $P(X = k + t \mid X \geqslant k) = P(X = t)$, we could get that

$$P(X = k + t + i \mid X \geqslant k) = P(X = t + i)$$

sum of i from 1 to ∞, and we could get that

$$P(X \geqslant k + t \mid X \geqslant k) = P(X \geqslant t)$$

7.1

Since F be the cdf of a discrete distribution that has the memoryless property as shown above, we could get that

$$\begin{aligned} 1 - F(h-1) &= P(X \geqslant h) \\ 1 - F(t+h-1) &= P(X \geqslant t + h) \\ 1 - F(t-1) &= P(X \geqslant t). \end{aligned}$$

And

$$P(X \geqslant h + t \mid X \geqslant h) = P(X \geqslant t)$$

By the definition, we have the conditional probability is

$$P(X \geqslant h + t \mid X \geqslant h) = \frac{P((X \geqslant h + t) \cap (X \geqslant h))}{P(X \geqslant h)}$$

Thus, we could get that

$$P(X \geqslant t) = \frac{P(X \geqslant h + t)}{P(X \geqslant h)}$$

Which means that

$$1 - F(h-1) = \frac{1 - F(t+h-1)}{1 - F(t-1)}$$

7.2

Since $l(x) = \log(1 - F(x-1))$, we have that

$$l(h) + l(t) = \log((1 - F(h-1))(1 - F(t-1))).$$

According to 7.1, we have that

$$(1 - F(h-1))(1 - F(t-1)) = 1 - F(t+h-1).$$

So

$$l(h) + l(t) = \log(1 - F(t+h-1)) = l(h+t)$$

7.3

For a integer t>0, we have that t=(t-1)+1. According to 7.2, we have that

$$l(t) = l(t-1) + l(1)$$

The rest can be done in the same manner, so we could get that

$$l(t) = l(t-1) + l(1) = l(t-2) + 2l(1) = tl(1)$$

Thus, we get that

$$l(t) = tl(1)$$

7.4

According to 7.1 we have that

$$P(X \geqslant t) = \frac{P(X \geqslant h+t)}{P(X \geqslant h)}.$$

Let t=0, we could get that $P(X \geqslant 0) = 1$. Let R(x)=1-F(x-1), so we have R(a+b)=R(a)R(b). Let R(1)=1-p, we could get that

$$R(n) = R(1)^n = (1-p)^n,\ R(n+1) = R(1)^{n+1} = (1-p)^{n+1}$$

So we could get that

$$P(X = n) = R(n) - R(n+1) = (1-p)^n p.$$

Which shows that X follows the geometric distribution and F must be the cdf of a geometric distribution.

8

If Y follows the normal distribution with mean 0 and variance 2, by the theorem, we could get the m.g.f. of Y is

$$M_Y(t) = e^{t^2},\ for\ -\infty < t < \infty.$$

So we have that the m.g.f. of X is the same as Y, by the theorem, the p.d.f. of X should be the same as Y. So we get the p.d.f. of X is

$$f_X(x) = \frac{1}{2\sqrt{\pi}} exp(-\frac{x^2}{4})$$

9

Since X follows the lognormal distribution with parameters μ and σ^2, let Y=log(X), so we could get the p.d.f. of Y is

$$f_Y(y) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right)$$

Let $Z = \frac{1}{X} = e^{-Y}$, so we could get the p.d.f. of Z is

$$f_Z(z) = \frac{1}{z}\frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{1}{2}\left(\frac{-\log(z)-\mu}{\sigma}\right)^2\right)$$

Thus, the p.d.f. of $\frac{1}{X}$ is

$$f_{\frac{1}{X}}\left(\frac{1}{x}\right) = \frac{x}{\sigma\sqrt{2\pi}} exp\left(-\frac{1}{2}\left(\frac{\log(\frac{1}{x})+\mu}{\sigma}\right)^2\right)$$

10

Since that X, Y are independent and each has the standard normal distribution with parameters $0, 1$, we could get the joint p.d.f. of X and Y is

$$f_{X,Y}(x, y) = f_X(x)f_Y(y) = \frac{1}{2\pi} exp\left(-\frac{1}{2}\left(x^2 + y^2\right)\right)$$

Let $Z = \frac{X}{Y},\ W = Y$, so we could get the Jacobi is

$$\mathbf{J} = det\begin{bmatrix} w & z \\ 0 & 1 \end{bmatrix} = w$$

So the joint c.d.f. of Z and W is

$$f(z, w) = \frac{1}{2\pi}\left|w\right| exp\left(-\frac{w^2(1 + z^2)}{2}\right)$$

Then the marginal p.d.f. of Z is

$$f_Z(z) = \int_{-\infty}^{\infty} f_{Z,W}(z, w)\, dw = \frac{1}{2\pi}\int_0^{\infty} exp\left(-\frac{w^2}{2}(1 + z^2)\right) dw^2 = \frac{1}{\pi(1 + z^2)}$$

Which shows that $Z = \frac{X}{Y}$ follows Cauchy distribution.

11

11.1

Since a random sample of size n is to be taken from the normal distribution with mean μ and standard deviation 2, we could get the p.d.f. of this normal distribution is

$$f_X(x) = \frac{1}{2\sqrt{2\pi}} exp\left(-\frac{(x - \mu)^2}{8}\right),\ for\ -\infty < x < \infty$$

By the definition, we could get that

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Since $(X_i)_{i=1}^n$ are i.i.d. and X_i follows $N(\mu, 4)$, by the corollary, we have that their sample mean $\overline{X}_n$ follows normal distribution with parameters μ and $\frac{4}{n}$.

So the p.d.f. of $Y = \overline{X}_n$ is

$$f_Y(y) = \frac{\sqrt{n}}{2\sqrt{2\pi}} exp\left(-\frac{n(y - \mu)^2}{8}\right),\ for\ -\infty < y < \infty$$

According to the table, since n is an integer, to satisfy

$$\mathbf{P}(|Y - \mu| < 0.1) \geqslant 0.9$$

Which means that

$$\int_{-\frac{\sqrt{n}}{20}}^{\frac{\sqrt{n}}{20}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}l^2}\, dl \geqslant 0.9$$

Searching for table, we could get that $\frac{\sqrt{n}}{20} \geqslant 1.65$, so the smallest n is 1089.

11.2

According to 11.1, for $\mu = 0.2$, n=20 and σ is unknown the p.d.f. of $Y = \overline{X}_n$ is

$$f_Y(y) = \frac{\sqrt{20}}{\sigma\sqrt{2\pi}} exp\left(-\frac{20(y-0.2)^2}{2\sigma^2}\right),\ for\ -\infty < y < \infty$$

To satisfy

$$\mathbf{P}(Y \leqslant 0.15) < 0.02$$

There must be

$$\int_{0.15}^{0.25} f_Y(y)\,dy \geqslant 0.96$$

Which means that

$$\int_{-\frac{1}{\sigma\sqrt{20}}}^{\frac{1}{\sigma\sqrt{20}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}l^2}\,dl \geqslant 0.96$$

Searching for table, we could get that $\frac{1}{\sigma\sqrt{20}} \geqslant 2.06$, so

$$\sigma \leqslant 0.1085$$

So σ can be 0.1085.

12

12.1

Since X follows Possion distribution with parameters $\alpha > 2$ and $\beta > 0$, we could have the p.d.f. of X is

$$f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x},\ for\ x > 0,$$

zero elsewhere.
By the definition, we have that

$$\mathbf{E}\left(\frac{1}{X}\right) = \int_0^\infty \frac{1}{x}\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}\,dx = \frac{\beta\Gamma(\alpha-1)}{\Gamma(\alpha)} = \frac{\beta}{\alpha-1}$$

12.2

By the definition, we have that

$$\mathbf{E}\left(\frac{1}{X^2}\right) = \int_0^\infty \frac{1}{x^2}\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}\,dx = \frac{\beta^2\Gamma(\alpha-2)}{\Gamma(\alpha)} = \frac{\beta^2}{(\alpha-1)(\alpha-2)}$$

By the theorem, we have that

$$\mathbf{Var}\left(\frac{1}{X^2}\right) = \mathbf{E}\left(\frac{1}{X^2}\right) - \mathbf{E}\left(\frac{1}{X}\right)^2 = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}$$

a