Homework 3

1. Suppose that I make an average of 3 mistakes per class. What is the PMF of X, the number of mistakes I will make next class? What is the probability that I will make at least one mistake?

2. A transmitter sends out either a 1 with probability p or a 0 with probability $1-p$, independently of earlier transmissions. If the number of transmissions within a given time interval has a Poisson PMF with parameter λ, show that the number of 1's transmitted in that same time interval has a Poisson PMF with parameter $p\lambda$.

3. A particular professor is known for his arbitrary grading policies. Each paper receives a grade from the set $\{A, A-, B+, B, B-, C+\}$ with equal probability, independent of other papers. How many papers do you expect to hand in before you receive each possible grade at least once?

4. Suppose that X is a normal random variable with mean 5. If $P(X > 9) = 0.2$, approximately what is $var(X)$?

5. Suppose that the height (in inches) of a 25-year old male living in Baltimore is a normal RV with mean 71 and variance 6.25. What percentage of 25-year old men in Baltimore are over 6 feet tall? What percentage of Baltimore 25-yr-old men who are taller than 6 feet are taller than 6 feet, 6 inches?

6. Show that the exponential distribution is memoryless. That is, show that $P(X > s + t \mid X > t) = P(X > s)$.

7. Show that Beta distribution family is an exponential family when:

   (a) α is a known constant and β is the only unknown parameter;

   (b) α is the only unknown parameter and β is constant;

   (c) both α and β are unknown parameters.

Statistical Theory Homework 3

Bohao Tang

September 13, 2017

1. It's reasonable to suppose that the number of mistakes you make in each class X is independent and follow a Poisson distribution $Poi(\lambda)$, then from the question we get that $\lambda = \mathbb{E}[X] = 3$. Therefore the PMF of X is:

$$\mathbf{P}(X = n) = \frac{3^n}{n!}e^{-3}$$

where $n \in \mathbb{N}$ ($0 \in \mathbb{N}$). And we have:

$$\mathbf{P}(X \geq 1) = 1 - \mathbf{P}(X = 0) = 1 - e^{-3}$$

2. Denote X to be the number of transmissions in this interval and Y be the number of 1's transmitted. Then we have:

$$\begin{aligned}
\mathbf{P}(Y = k) &= \sum_{n=0}^{\infty} \mathbf{P}(Y = k; X = n) = \sum_{n=0}^{\infty} \mathbf{P}(Y = k | X = n)\mathbf{P}(X = n) && (1)\\
&= \sum_{n=k}^{\infty} \mathbf{P}(Y = k | X = n)\mathbf{P}(X = n) = \sum_{n=k}^{\infty} \binom{n}{k} p^k (1-p)^{n-k} \frac{\lambda^n}{n!} e^{-\lambda} && (2)\\
&= \frac{e^{-\lambda}p^k\lambda^k}{k!} \sum_{n=k}^{\infty} \frac{[(1-p)\lambda]^{n-k}}{(n-k)!} = \frac{e^{-\lambda}p^k\lambda^k}{k!} \cdot e^{\lambda - p\lambda} && (3)\\
&= \frac{(p\lambda)^k}{k!} e^{-p\lambda} && (4)
\end{aligned}$$

Therefore Y $\sim Poi(p\lambda)$.

3. Let X be the number of papers you hand in before you receive each possible grade at least once. Actually I think there are two interpretation of the word "before". One is that you hand in your X^{th} paper, and after grading, you firstly get every grade. And the other is that right after you send your X^{th} paper (that is when you send your $(X+1)^{th}$ paper), you firstly get every grade.

In this solution, I think the first interpretation is more meaningfull, although they are nearly the same (if you use the second interpretation, plus my answer with 1).

We then easy to see that $X = X_1 + X_2 + X_3 + X_4 + X_5 + X_6$, where X_i is the paper you need to get a new type grade after you already get $i - 1$ types of grades. Also, X_i will follow a geometric distribution with success probability $\frac{6-i+1}{6}$. Because every grading of paper is independent, to get a new type of grade, you are just repeat doing iid experiments until success.

Therefore we have $\mathbb{E}[X] = \sum_{i=1}^{6} \mathbb{E}[X_i] = \sum_{i=1}^{6} \frac{6}{7-i} = \frac{147}{10}$.

4. Suppose $var(X) = \sigma^2$, then $(X-5)/\sigma$ follows the standard normal distribution. Then $\mathbf{P}(X > 9) = 0.2 \Rightarrow \mathbf{P}(\frac{X-5}{\sigma} > \frac{4}{\sigma}) = 0.2$, call R function "qnorm(0.8)", we can get that $\frac{4}{\sigma} = 0.8416212$. Therefore $var(X) = 22.58846$.

5. We use R function "pnorm" to do the calculation. Here suppose H be the random variable of height:

$$\begin{aligned}
\mathbf{P}(H > 6') &= 1 - \mathbf{P}(H \leq 72") && (5)\\
&= 1 - \mathbf{P}(\frac{H-71}{\sqrt{6.25}} \leq \frac{72-71}{\sqrt{6.25}}) && (6)\\
&= 1 - \text{pnorm}(\frac{72-71}{\sqrt{6.25}}) && (7)\\
&= 0.3445783\cdots \doteq 34.458\% && (8)\\
\mathbf{P}(H > 6'6"|H > 6') &= \frac{\mathbf{P}(H > 78; H > 72)}{\mathbf{P}(H > 72)} && (9)\\
&= \frac{1 - \text{pnorm}(\frac{78-71}{\sqrt{6.25}})}{1 - \text{pnorm}(\frac{72-71}{\sqrt{6.25}})} = 0.00741525\ldots && (10)\\
&= 0.7415\% && (11)
\end{aligned}$$

6. Suppose X follows exponential distribution $exp(\lambda)$, then we have $\mathbf{P}(X > s) = e^{-\lambda s}$, for all $s \geq 0$. Therefore:

$$\begin{aligned}
\mathbf{P}(X > s+t|X > t) &= \frac{\mathbf{P}(X > s+t; X > t)}{\mathbf{P}(X > t)} = \frac{\mathbf{P}(X > s+t)}{\mathbf{P}(X > t)} && (12)\\
&= \frac{e^{-s-t}}{e^{-t}} = e^{-s} = \mathbf{P}(X > s) && (13)
\end{aligned}$$

7. Suppose X follows a Beta distribution $Be(\alpha, \beta)$. Then its pdf is:

$$f(x; \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}$$

(a) Given α a known constant, we have:

$$f(x; \beta) = \frac{1}{B(\alpha, \beta)} \cdot x^{\alpha-1}\mathbb{1}_{0\leq x\leq 1} \cdot exp\{(\beta-1)\ln(1-x)\}$$

Therefore it is of exponential family with $h(\theta) = \frac{1}{B(\alpha,\beta)}$; $c(x) = x^{\alpha-1}\mathbb{1}_{0\leq x\leq 1}$; $\omega(\theta) = (\beta-1)$; $t(x) = \ln(1-x)$.

(b) Given β a known constant, we have:

$$f(x; \alpha) = \frac{1}{B(\alpha, \beta)} \cdot (1-x)^{\beta-1}\mathbb{1}_{0\leq x\leq 1} \cdot exp\{(\alpha-1)\ln(x)\}$$

Therefore it is of exponential family with $h(\theta) = \frac{1}{B(\alpha,\beta)}$; $c(x) = (1-x)^{\beta-1}\mathbb{1}_{0\leq x\leq 1}$; $\omega(\theta) = (\alpha-1)$; $t(x) = \ln(x)$.

(c) In general, we have:

$$f(x; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} \cdot \mathbb{1}_{0\leq x\leq 1} \cdot exp\{(\alpha-1)\ln(x) + (\beta-1)\ln(1-x)\}$$

Therefore it is of exponential family with $h(\theta) = \frac{1}{B(\alpha,\beta)}$; $c(x) = \mathbb{1}_{0\leq x\leq 1}$; $\omega_1(\theta) = (\alpha-1)$; $\omega_2(\theta) = (\beta-1)$; $t_1(x) = \ln(x)$; $t_2(x) = \ln(1-x)$.