SI140 Probability & Mathematical Statistics

Homework 6

陈昱聪

Chen Yucong ><E<>N

Student ID: 2019533079

Email: chenyc@shanghaitech.edu.cn

School of Information Science and Technology

◎ Group#2 (TA: 曾理)

Problem 4.20

Solution:

(a) Since we can say that $V + W \equiv X + Y$, we have:

$$E(V) + E(W) = E(V + W) = E(X + Y) = E(X) + E(Y) = \frac{n}{2} + \frac{n+1}{2} = n + \frac{1}{2}$$

(b) Since $p = \frac{1}{2}$, we can find the symmetry between the number of successes and the number of failures.

That is $p = q = \frac{1}{2}$, so $n - X \sim \text{Bin}(n, q) = \text{Bin}(n, \frac{1}{2})$, $n + 1 - Y \sim \text{Bin}(n + 1, q) = \text{Bin}(n + 1, \frac{1}{2})$.

That means the distributions of X and $n - X$ are identical, and so are the distributions of Y and $n + 1 - Y$, while X and Y are independent.

So we have $P(X < Y) = P(n - X < n + 1 - Y)$.

(c) Using (b) we know

$$\begin{aligned}
P(X < Y) &= P(n - X < n + 1 - Y) \\
&= P(Y < X + 1) \\
&= P(Y \leqslant X) \quad \text{(Since X and Y are integer-valued)} \\
&= 1 - P(X < Y)
\end{aligned}$$

So that $P(X < Y) = \frac{1}{2}$

Problem 4.25

Solution:

(a) In the k^{th} trial, $P(X_k = Y_k = 1) = p_1p_2$. So that $n \sim \text{FS}(p_1p_2)$, which means $E(n) = \frac{1}{p_1p_2}$.

(b) That is, in k^{th} trail $X_kY_k = 1$, we get $P(X_kY_k = 1) = (1-p_1)(1-p_2)$. So that $n \sim \text{FS}((1-p_1)(1-p_2))$, which means $E(n) = \frac{1}{(1-p_1)(1-p_2)}$.

(c) This case is that there are several times (could be 0 to ∞) where they are both failed. So

$$\begin{aligned}
P(\text{Their first successes are simultaneous}) &= p_1p_2\sum_{k=0}^{\infty}[(1-p_1)(1-p_2)]^k \\
&= p_1^2 \lim_{k\to\infty} \frac{1-(1-p_1)^{2k}}{1-(1-p_1)^2} \\
&= \frac{p_1}{2-p_1}
\end{aligned}$$

By using symmetry, the probability of Nick's preceding equals to the probability of Penny's preceding. So we have

$$P(\text{Nick's preceding}) = \frac{1-\frac{p_1}{2-p_1}}{2} = \frac{1-p_1}{2-p_1}$$

Problem 4.27

Solution:

(a)

$$\begin{aligned}
E(Xg(X)) = \sum_{k=0}^{\infty} kg(k)P(X=k) &= \sum_{k=0}^{\infty} kg(k)\frac{e^{-\lambda}\lambda^k}{k!} \\
&= \sum_{k=1}^{\infty} kg(k)\frac{e^{-\lambda}\lambda^k}{k!} \\
&= \lambda\sum_{k=1}^{\infty} g(k)\frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!} \\
&= \lambda\sum_{k=0}^{\infty} g(k+1)\frac{e^{-\lambda}\lambda^k}{k!} \\
&= \lambda\sum_{k=0}^{\infty} g(k+1)P(X=k) \\
&= \lambda E(g(X+1))
\end{aligned}$$

(b) Let $g(X) = X^2$, by using $E(X) = \lambda$ and $E(X^2) - E^2(X) = \lambda$, we know that

$$\begin{aligned}
E(X^3) &= E(X \cdot X^2) \\
&= \lambda E(X^2 + 2X + 1) \\
&= \lambda E(X^2) + 2\lambda E(X) + \lambda \\
&= \lambda(\lambda + E^2(X)) + 2\lambda E(X) + \lambda \\
&= \lambda(\lambda + \lambda^2) + 2\lambda^2 + \lambda \\
&= \lambda^3 + 3\lambda^2 + \lambda
\end{aligned}$$

Problem 4.38

Solution:

Let

$$I_i = \begin{cases} 1 & \text{(The } i^{\text{th}} \text{ type of toys is collected.)} \\ \\ 0 & \text{(The } i^{\text{th}} \text{ type of toys is not collected.)} \end{cases}$$

So the expected number of distinct toy types can be caculated as follows:

$$E\left(\sum_{i=1}^{n} I_i\right) = \sum_{i=1}^{n} E(I_i) = nE(I_i) = nE(I_1)$$

By using Bose-Einstein Counting, We know

$$\begin{aligned} E(I_1) &= P(\text{The } 1^{\text{st}} \text{ type of toys is collected.}) \\ &= 1 - P(\text{The } 1^{\text{st}} \text{ type of toys is not collected.}) \\ &= 1 - \frac{\binom{n+t-2}{n-2}}{\binom{n+t-1}{n-1}} \\ &= \frac{t}{n+t-1} \end{aligned}$$

So the expectation is:

$$\frac{nt}{n+t-1}$$

Problem 4.44

Solution:

(a) It means the expected number of pairs of both w in those n objects. So

$$E\binom{X}{2} = \binom{n}{2}\frac{w}{w+b}\frac{w-1}{w+b-1}$$

(b) By using (a),

$$E\binom{X}{2} = \frac{1}{2}E(X(X-1)) = \frac{1}{2}(E(X^2)-E(X)) = \binom{n}{2}\frac{w}{w+b}\frac{w-1}{w+b-1}$$

$$\Rightarrow E(X^2) - X(X) = n(n-1)p\frac{w-1}{N-1}, \quad \text{while } E(X) = np$$

$$\begin{aligned} V(X) &= E(X^2) - E^2(X) \\ &= n(n-1)p\frac{w-1}{N-1} + np - n^2p^2 \\ &= \frac{N-n}{N-1}npq \end{aligned}$$

Problem 4.48

Solution:

Let

$$I_i = \begin{cases} 1 & (\text{The } i^{\text{th}} \text{ toss is diffetent from the } (i-1)^{\text{th}}.) \\ 0 & (\text{The } i^{\text{th}} \text{ toss is the same as the } (i-1)^{\text{th}}.) \end{cases}$$

Where $i = 2, 3, 4, \ldots, n$.
So the expected #runs is:

$$E(1+\sum_{i=2}^{n} I_i) = 1 + \sum_{i=2}^{n} E(I_i) = 1 + \sum_{i=2}^{n} P(I_i = 1) = 1 + 2(n-1)p(1-p)$$

Problem 4.54

Solution:

(a) With respect to CDF of the birthday problem, $P(X \leqslant 23) = 50.7\%$, so that $P(X \leqslant 24) > \frac{1}{2}$ since with the increase of X, CDF is absolutely monotonous. And we know $P(X \leqslant 22) = \frac{1}{2}$, $P(X \geqslant 23) = 1 - P(X < 22) > \frac{1}{2}$. So 23 is the median while 22 and 24 are not. And since the CDF is monotonous so we know if we have multiple medians, they should be in together. So the 23 is the unique median.

(b) From the description, we found that

$$E(I_j) = P(I_j = 1) = P(X \geqslant j) = p_j$$

And after basic analysis we found the expression of p_j as defined in this Question. Using the equation we found above,

$$E(X) = \sum_{j=1}^{366} E(I_J) = \sum_{j=1}^{366} p_j$$

(c) $E(X) = 24.6166$ by using Python code. (See the Appendix)

(d) Since $i < j$ in that equation so that if $I_j = 1$ then $I_i = 1$ for sure. Now that we want $I_i I_j = 1$, that means $I_j = 1$.

$$E(X^2) = \sum_{j=1}^{366} E(I_j^2) + 2\sum_{j=2}^{366}\sum_{i=1}^{j-1}(1 \cdot I_j) = \sum_{j=1}^{366} E(I_j) + 2\sum_{j=2}^{366}((j-1)\cdot I_j) = \sum_{j=1}^{366} p_j + 2\sum_{j=2}^{366}(j-1)p_j$$

$$V(X) = E(X^2) - E^2(X) = \sum_{j=1}^{366} p_j + 2\sum_{j=2}^{366}(j-1)p_j - \left(\sum_{j=1}^{366} p_j\right)^2 = 148.6403 \quad \text{(See the Appendix)}$$

APPENDIX (Code With Python):

```python
Ex = 2
Vx = 1
p = [1]*366

for j in range(2, 366):
    p[j] = p[j-1]*(1-(j - 1)/365)
    Ex += p[j]
    Vx += j*p[j]
Vx = Ex - Ex**2 + 2*Vx

print("Ex = ", Ex, "Vx = ", Vx)
```

Problem 4.61

Solution:

(a) Biased.

$$\begin{aligned}
E(T)-\theta &= \sum_{k=0}^{\infty} e^{-3k}\frac{e^{-\lambda}\lambda^k}{k!} - e^{-3\lambda} \\
&= e^{-\lambda}\sum_{k=0}^{\infty}\frac{e^{-3k}\lambda^k}{k!} - e^{-3\lambda} \\
&= e^{-\lambda}\sum_{k=0}^{\infty}\frac{(e^{-3}\lambda)^k}{k!} - e^{-3\lambda} \\
&= e^{-\lambda}e^{(e^{-3}\lambda)} - e^{-3\lambda} \neq 0
\end{aligned}$$

(b)

$$\begin{aligned}
E(g(X))-\theta &= \sum_{k=0}^{\infty}(-2)^k\frac{e^{-\lambda}\lambda^k}{k!} - e^{-3\lambda} \\
&= e^{-\lambda}\sum_{k=0}^{\infty}\frac{(-2)^k\lambda^k}{k!} - e^{-3\lambda} \\
&= e^{-\lambda}\sum_{k=0}^{\infty}\frac{(-2\lambda)^k}{k!} - e^{-3\lambda} \\
&= e^{-\lambda}e^{-2\lambda} - e^{-3\lambda} = 0
\end{aligned}$$

(c) Beacuse $(-2)^X$ is sometimes negative while θ is always positive. Let

$$h(X) = \begin{cases} 0\,, & \text{X is odd} \\ g(X)\,, & \text{X is even} \end{cases}$$

Since θ is positive so using 0 to estimating is better in the part where $g(x)$ is negative.

So we get $|h(X)-\theta| \leqslant |g(X)-\theta|$

Problem 4.63

Solution:

(a) Let I_j be the indicator for whether the j^{th} guy picking his own name. Then,

$$E(X) = E\left(\sum_{j=1}^{n} I_j\right) = \sum_{j=1}^{n} E(I_J) = n \cdot \frac{1}{n} = 1$$

(b) There are $\binom{n}{2} = \frac{n(n-1)}{2}$ pairs, and for two guys we got the probability $p = \frac{1}{n(n-1)}$ that they pick each other's name. So the expectation is $\frac{n(n-1)}{2} \cdot \frac{1}{n(n-1)} = \frac{1}{2}$

(c) As we know the expectation is 1, we would construct the distribution as $X \sim \text{Pois}(1)$ with $\lambda = 1$. As $n \to \infty$, $P(X = 0) \to \frac{e^{-\lambda}\lambda^0}{0!} = \frac{1}{e}$

Problem 4.72

Proof:

Using probability to solve this problem. We can generate m arbitrary strings with length n to check if the probability of k-complete is positive. If so, we can say that there exists a k-complete set of size m. Let $N = \binom{n}{k}2^k$, and let A_j be the event that S contains the i^{textth} string as the question described. Then using the theorem we have:

$$P\left(\bigcup_{j=1}^{N} A_j^c\right) \leqslant \sum_{j=1}^{N} P(A_j^c) = \binom{n}{k} 2^k (1 - 2^{-k})^m < 1$$

$$\Rightarrow P\left(\bigcap_{j=1}^{N} A_j\right) > 0$$

□

Problem 4.83

Solution:

(a) After exploration phase, the sum of the ranks is $k \cdot \frac{n+1}{2}$ since the expectation of a rank with randomly choice from n dishes is the population mean that $\frac{n+1}{2}$.

In exploitation phase we always order the best dish for $m-k$ times. Let Y be the sum after all. So after all we have the expected sum:

$$E(Y) = k\,\frac{n+1}{2} + (m-k)E(X)$$

(b) The total number of ways in odering is $\binom{n}{k}$. The number of odering ways while $X=j$ is $\binom{k-1}{j-1}$ since once we have the largest rank, the rest that $k-1$ ranks could only be chosen in which is less than the largest one. So we have the PMF

$$P(X=j) = \frac{\binom{k-1}{j-1}}{\binom{n}{k}}$$

(c) We know that $X \in \{k, k+1, \ldots, n\}$ since it is the max in k orders.

$$\begin{aligned}
E(X) &= \sum_{j=k}^{n} jP(X=j) = \sum_{j=k}^{n} \frac{\binom{k-1}{j-1}}{\binom{n}{k}} = \sum_{j=k}^{n} j\,\frac{\frac{(j-1)!}{(j-k)!(k-1)!}}{\frac{n!}{(n-k)!k!}} \\
&= \frac{(k+1)!(n-k)!}{n!}\sum_{j=k}^{n}\frac{j!}{(j-k)!k!} = \frac{(k+1)!(n-k)!}{n!}\sum_{j=k}^{n}\binom{j}{k} \\
&= \frac{(k+1)!(n-k)!}{n!}\binom{n+1}{k+1} \quad \text{(Using the hockey stick identity)} \\
&= \frac{k(n+1)}{k+1}
\end{aligned}$$

(d)

$$E(Y) = k\,\frac{n+1}{2} + (m-k)E(X) = k\,\frac{n+1}{2} + (m-k)\frac{k(n+1)}{k+1}$$

Get the maximum value where the derivative is zero:

$$\left.\frac{\mathrm{d}E(Y)}{\mathrm{d}k}\right|_{k_0} = \frac{n+1}{2} - \frac{(2k_0 + k_0^2 - m)(1+n)}{(1+k_0)^2} = 0$$

$$\Rightarrow k_0 = \sqrt{2(m+1)} - 1$$