Tutorial 5

1. A RV X has symmetric probability mass function (PMF) such that $P_X(k) = P_X(-k)$, $k = \cdots, -1, 0, 1, \cdots$. Prove that all odd order moments are equal to zero, i.e., $\mathbb{E}\{X^n\} = 0$ for all odd numbers of n.

2. Use the definition of $\mathbb{E}\{X\}$ in (2.19) to compute $\mathbb{E}\{X\}$ of the binomial random variable with parameters n and p, whose PMF is:

$$p(r) = C(n,r)p^r(1-p)^{n-r}, \quad 0 \leq r \leq n$$

3. Use the definition of the moment generating function $\phi(t)$ in (2.28) to compute $\mathbb{E}\{X\}$ of the exponential random variable with parameter λ, whose probability density function (PDF) is:

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

4. Compute $\mathbb{E}\{X\}$ of the geometric random variable with parameter p, whose PMF is:

$$p(r) = (1-p)^{r-1}p, \quad 1 \leq r < \infty$$

5. Consider the experiment of rolling a dice where the outcome is the face number. Suppose the probability of obtaining a "1", "2", and "3" is the same, while the probability of obtaining a "4", "5", and "6" is the same. However, a "5" is twice as likely to be observed as a "1". What is the expected value of the face number?

6. Consider a data sequence only with letters "A", "B", "C", "D", and the current encoding scheme adopts a 2-bit code, 00, 01, 10, 11, respectively, i.e., each letter needs 2 bits for storage.

   Suppose we know that the probabilities of occurrence of "A", "B", "C" and "D" are 7/8, 1/16, 1/32, and 1/32, respectively. To utilize this information, we investigate encoding "A", "B", "C" and "D", by 0, 10, 110, and 111. Compute the expected number of bits per letter based on this strategy.

Solution

1.

Analogous to (2.21), we have:

$$\begin{aligned}
\mathbb{E}\{X^n\} &= \sum_{k=-\infty}^{\infty} k^n P_X(k) \\
&= \sum_{k=-\infty}^{-1} k^n P_X(k) + \sum_{k=1}^{\infty} k^n P_X(k) \\
&= \sum_{l=1}^{\infty} (-l)^n P_X(-l) + \sum_{k=1}^{\infty} k^n P_X(k) \\
&= \sum_{l=1}^{\infty} (-l)^n P_X(l) + \sum_{k=1}^{\infty} k^n P_X(k) \\
&= \sum_{k=1}^{\infty} [(-k)^n + k^n] P_X(k)
\end{aligned}$$

For odd n, $(-k)^n + k^n = 0$, and thus $\mathbb{E}\{X^n\} = 0$.

2.

$$
\begin{aligned}
\mathbb{E}\{X\} = \sum_{r=0}^{n} r p(r) &= \sum_{r=0}^{n} r \binom{n}{r} p^r (1-p)^{n-r} \\
&= \sum_{r=1}^{n} \frac{rn!}{(n-r)!r!} p^r (1-p)^{n-r} \\
&= \sum_{r=1}^{n} \frac{n!}{(n-r)!(r-1)!} p^r (1-p)^{n-r} \\
&= np \sum_{r=1}^{n} \frac{(n-1)!}{(n-r)!(r-1)!} p^{r-1} (1-p)^{n-r} \\
&= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k} \\
&= np \left[p + (1-p)\right]^{n-1} = np
\end{aligned}
$$

3.

$$\phi(t) = \mathbb{E}\{e^{tX}\} = \int_{-\infty}^{\infty} e^{tx} p(x) dx = \int_{0}^{\infty} e^{tx} \lambda e^{-\lambda x} dx = \int_{0}^{\infty} \lambda e^{(t-\lambda)x} dx$$

$$= \lambda \frac{1}{t-\lambda} e^{(t-\lambda)x} \Big|_{0}^{\infty}$$

$$= \frac{\lambda}{\lambda - t}, \quad \lambda > t$$

Note that $\phi(t)$ is only defined for $\lambda > t$ but this does not affect our calculation because $\lambda > 0$ and we evaluate at $t = 0$.

$$\phi'(t) = \frac{\lambda}{(\lambda - t)^2} \Rightarrow \phi'(0) = \mathbb{E}\{X\} = \frac{\lambda}{(\lambda)^2} = \frac{1}{\lambda}$$

Note that when using the moment generating function, it is simpler than using (2.20) because there is no need to deal with the integration of $x\lambda e^{-\lambda x}$ which requires integrating by parts.

4.

We can use (2.19):

$$
\begin{aligned}
\mathbb{E}\{X\} = \sum_{r=1}^{\infty} rp(r) &= \sum_{r=1}^{\infty} rp(1-p)^{r-1} \\
&= p\sum_{r=1}^{\infty} rq^{r-1}, \quad q = 1-p \\
&= p\sum_{r=1}^{\infty} \frac{d}{dq} q^r = p\frac{d}{dq}\sum_{r=1}^{\infty} q^r \\
&= p\frac{d}{dq}\frac{q}{1-q} = p\frac{(1-q)-q(-1)}{(1-q)^2} = p\frac{1}{(1-q)^2} = \frac{1}{p}
\end{aligned}
$$

That is, the expected number of independent trials we need to perform until the first success equals the reciprocal of the probability of a success in a trial, e.g., if $p = 0.1$, then on average it takes $1/p = 10$ trials for a success.

We can also use (2.31):

$$\phi(t) = \sum_{r=1}^{\infty} e^{tr} p(1-p)^{r-1} = pe^t \sum_{r=1}^{\infty} [e^t(1-p)]^{r-1} = pe^t \sum_{k=0}^{\infty} [e^t(1-p)]^k$$

$$= \frac{pe^t}{1-e^t(1-p)}, \quad |e^t(1-p)| < 1$$

$$\phi'(t) = \frac{[1-e^t(1-p)]pe^t - (pe^t)[-e^t(1-p)]}{[1-e^t(1-p)]^2} = \frac{[1-e^t(1-p)]pe^t + p(1-p)e^{2t}}{[1-e^t(1-p)]^2}$$

Hence

$$\Rightarrow \phi'(0) = \mathbb{E}\{X\} = \frac{[1-(1-p)]p + p(1-p)}{[1-(1-p)]^2} = \frac{1}{p}$$

5.

Assigning the random variable X as the face number, we have $1 \le X \le 6$. Let the probability of getting "1" be p. Then:

$$p(1) = p(2) = p(3) = p \quad \text{and} \quad p(4) = p(5) = p(6) = 2p$$

As the sum of all PMFs is 1, we easily obtain $p = 1/9$. The expected value of the face number is thus:

$$\mathbb{E}\{X\} = \frac{(1+2+3)+2(4+5+6)}{9} = 4$$

6.

Assigning the random variable X as the codelength in terms of bit number, we have $1 \le X \le 3$. That is, for the sample space {A,B,C,D}, we have X(A)=1, X(B)=2, X(C)=X(D)=3.

Based on the given probability information, the PMF for X is $p(1) = 7/8$, $p(2) = 1/16$, and $p(3) = 1/32 + 1/32 = 1/16$. As a result,

$$\mathbb{E}\{X\} = 1 \cdot \frac{7}{8} + 2 \cdot \frac{1}{16} + 3 \cdot \frac{1}{16} = 1.1875$$

This means that the average number of bits for a letter is reduced from 2 to 1.1875.

Note that the idea of assigning shorter code words to the letters that occur more often is referred to as Huffman coding.