Statistics I, Fall 2012
Suggested Solution for Homework 06

Ling-Chieh Kung
Department of Information Management
National Taiwan University

1. We have

$$\begin{aligned}\mathbb{E}\big[X^2\big] &= \int_0^\infty x^2 \lambda e^{-\lambda x} dx \\ &= \lambda\left[x^2\left(-\frac{1}{\lambda}e^{-\lambda x}\right)\Big|_0^\infty - \int_0^\infty -\frac{1}{\lambda}e^{-\lambda x} \cdot 2x dx\right] = \lambda\left(0 + \frac{2}{\lambda}\int_0^\infty x e^{-\lambda x} dx\right),\end{aligned}$$

where $x^2(-\frac{1}{\lambda}e^{-\lambda x})|_0^\infty$ equals 0 because

$$\lim_{x\to\infty} \frac{x^2}{e^{\lambda x}} = 0 \quad \forall \lambda > 0.$$

Now, note that $\int_0^\infty x e^{-\lambda x} dx$ has been evaluated (with another integration by parts) when we calculated $\mathbb{E}[X]$ during lectures:

$$\int_0^\infty x e^{-\lambda x} dx = \frac{1}{\lambda^2}.$$

So $\mathbb{E}[X^2] = \lambda(\frac{2}{\lambda})(\frac{1}{\lambda^2}) = \frac{2}{\lambda^2}$ and $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{2}{\lambda^2} - (\frac{1}{\lambda})^2 = \frac{1}{\lambda^2}$.

2. We have

$$\begin{aligned}\mathbb{E}\big[(X_1 - 2X_2)^2\big] &= \mathbb{E}\big[X_1^2 - 4X_1X_2 + 4X_2^2\big] = \mathbb{E}\big[X_1^2\big] - 4\mathbb{E}\big[X_1X_2\big] + 4\mathbb{E}\big[X_2^2\big] \\ &= \mathbb{E}\big[X_1^2\big] - 4\mathbb{E}\big[X_1\big]\mathbb{E}\big[X_2\big] + 4\mathbb{E}\big[X_2^2\big] = \frac{1}{3} - 1 + \frac{4}{3} = \frac{2}{3},\end{aligned}$$

where $\mathbb{E}[X_1^2] = \mathbb{E}[X_2^2] = \int_0^1 x^2 \cdot 1 dx = \frac{1}{3}$ and $\mathbb{E}[X_1] = \mathbb{E}[X_2] = \int_0^1 x \cdot 1 dx = \frac{1}{2}$.

3. Let X be the number of letters contained in the selected word, we may construct the probability mass function for X as in Table 1.

x	1	2	3	4	6	7	8
$\Pr(X = x)$	$\frac{1}{13}$	$\frac{2}{13}$	$\frac{1}{13}$	$\frac{4}{13}$	$\frac{2}{13}$	$\frac{1}{13}$	$\frac{2}{13}$

Table 1: Probability mass function for Problem 3.

(a) $\Pr(X \geq 6) = \frac{2}{13} + \frac{1}{13} + \frac{2}{13} = \frac{5}{13} \approx 0.385$.

(b) $\mathbb{E}[X] = 1 \times \frac{1}{13} + 2 \times \frac{2}{13} + \cdots + 8 \times \frac{2}{13} = \frac{59}{13} \approx 4.538$.

(c) $\mathrm{Var}(X) = (1 - \frac{59}{13})^2 \times \frac{1}{13} + (2 - \frac{59}{13})^2 \times \frac{2}{13} + \cdots + (8 - \frac{59}{13})^2 \times \frac{2}{13} \approx 4.864$.

4. We have

$$\mathbb{E}[X] = \sum_{x<t} x \Pr(x) + \sum_{x \geq t} x \Pr(x) \geq 0 + \sum_{x \geq t} t \Pr(x) = t \Pr(X \geq t),$$

where the inequality holds because $x \geq 0$ in the first term and $x \geq t$ in the second term. Dividing both sides by t then yields $\Pr(X \geq t) \leq \frac{\mathbb{E}[X]}{t}$.

5. (a) We apply the separability of expectation and $\mathbb{E}[X_i] = \mu$ to obtain

$$\mathbb{E}\big[\overline{X}\big] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n X_i\right] = \frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i] = \frac{1}{n}\sum_{i=1}^n \mu = \frac{1}{n}(n\mu) = \mu.$$

(b) We apply the separability of variance for independent random variables and $\text{Var}(X_i) = \sigma^2$ to obtain

$$\text{Var}(\overline{X}) = \text{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\text{Var}(X_i) = \frac{1}{n^2}(n\sigma^2) = \frac{\sigma^2}{n}.$$

(c) By Chebyshev's theorem, we have

$$\Pr(|\overline{X} - \mu| > k\frac{\sigma}{\sqrt{n}}) \leq \frac{1}{k^2},$$

where $\frac{\sigma}{\sqrt{n}}$ is the standard deviation of $\overline{X}$. By setting $\epsilon = \frac{k\sigma}{\sqrt{n}}$, we get $k^2 = \frac{\sigma^2}{n\epsilon^2}$, which implies

$$\Pr(|\overline{X} - \mu| > \epsilon) \leq \frac{\delta^2}{n\epsilon^2}.$$

Therefore, because $\Pr(|\overline{X} - \mu| > \epsilon) \geq 0$ and $\lim_{n\to\infty} \frac{\delta^2}{n\epsilon^2} = 0$,

$$\lim_{n\to\infty} \Pr(|\overline{X} - \mu| > \epsilon) = 0$$

by the squeeze theorem.

6. (a) The frequency distribution is contained in Table 2 and the frequency polygon is illustrated in Figure 1.

Class	Observation	Theoretical
$[-20, -5)$	2	2.48
$[-5, 10)$	9	9.64
$[10, 25)$	32	29.17
$[25, 40)$	76	68.69
$[40, 55)$	122	125.90
$[55, 70)$	187	179.63
$[70, 85)$	184	199.50
$[85, 100)$	181	172.48
$[100, 115)$	110	116.08
$[115, 130)$	60	60.81
$[130, 145)$	28	24.80
$[145, 160)$	9	7.87

Table 2: Frequency distribution for Problem 6.

(b) According to the histogram, we may get a feeling that the data are normally distributed. In order to test this intuition, let's try to approximate this set of data by a normal distribution. In choosing the parameters of the normal distribution, it is natural to use the sample mean and sample standard deviation, which are 76.28 and 29.66, respectively.

Now, if the normal distribution provides a good approximation, in each class the observed frequency and the theoretical frequency derived from the normal distribution should be close to each other. For example, in the first class $[-20, -5)$, we have observed 2 occurrences. Suppose $X \sim \text{ND}(76.28, 29.66)$, we may calculate that $\Pr(-20 \leq X < -5) \approx 0.00248$ and thus among the 1000 observations, $1000 \times 0.00248 = 2.48$ occurrences should be observed in this class. As 2 and 2.48 seems to be close to each other, at least in the first class the normal approximation is not bad.

We may then repeat this process for all the classes to make a comparison. The theoretical number of observations are listed in Table 2 and the two frequency polygons, one for the true observations and one for the theoretical observations, are plotted in Figure 2. As these two frequency polygons are quite close to each other, we may conclude (with only intuitions) that the data follow a normal distribution with mean 76.28 and standard deviation 29.66.

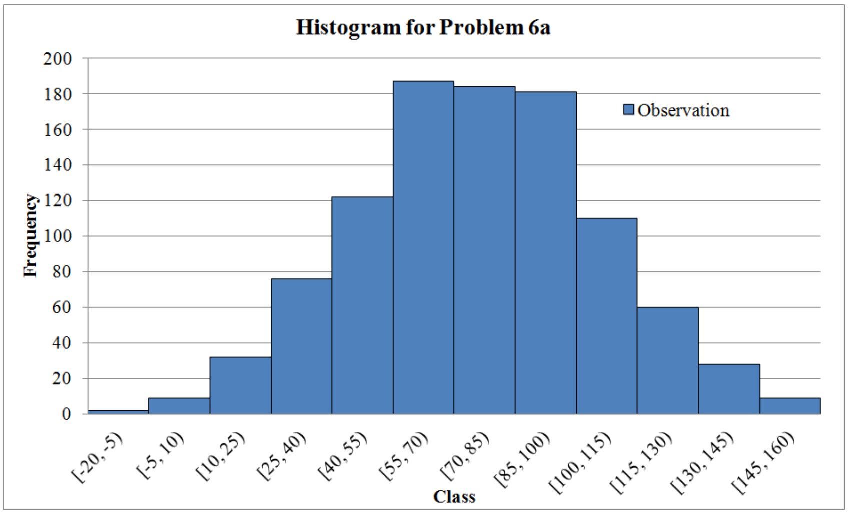

Figure 1: Histogram for Problem 6a.

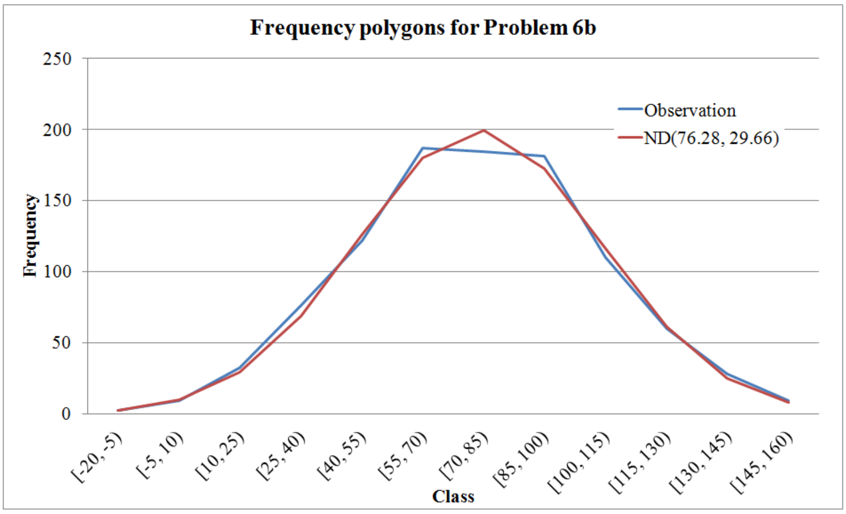

Figure 2: Frequency polygons for Problem 6b.

7. Let $X \sim \text{ND}(180, 32)$ and $Z \sim \text{ND}(0, 1)$.

   (a) By transforming into the standard normal distribution, we may try to find x such that

   $$\Pr(X > x) = \Pr\left(Z > \frac{x - 180}{32}\right) = 0.7.$$

   By looking at the standard normal probability table or using the MS Excel function NORMSINV(), we may find that $\Pr(Z > z) = 0.7$ if and only if $z = \frac{x-180}{32} = -0.5244$. It then follows that $x = 180 - 32 \times 0.5244 = 163.22$. You may also use the MS Excel function NORMINV() to find x directly. Nevertheless, you may need to know how to do transformation and use the probability table in the midterm exam.

   (b) Again,we may try to find x such that

   $$\Pr(X < x) = \Pr\left(Z > \frac{x - 180}{32}\right) = 0.2.$$

   As $\Pr(Z > z) = 0.2$ if and only if $z = \frac{x-180}{32} = 0.8416$, $x = 180 + 32 \times 0.8416 = 206.93$.

8. Let $\lambda = 6.4$ arrivals per minute be the arrival rate, $X \sim \text{Poi}(6.4)$ be the number of arrivals in one minute, and $Y \sim \text{Exp}(6.4)$ be the number of minutes between two arrivals.

   (a) The average interarrival time is $\mathbb{E}[Y] = \frac{1}{\lambda} = \frac{1}{6.4}$ minutes or $60 \times \frac{1}{6.4} = 9.375$ seconds.

   (b) $\Pr(X \geq 10) \approx 0.114$.
   Note. You may use the MS Excel function POISSON(9, 6.4, 1) to find $\Pr(X \leq 9)$.

   (c) $\Pr(Y \geq \frac{25}{60}) \approx 0.069$.
   Note. You may use the MS Excel function EXPONDIST($\frac{25}{60}$, 6.4, 1) to find $\Pr(Y < \frac{25}{60})$.

   (d) $\Pr(Y \leq \frac{5}{60}) \approx 0.413$.
   Note. You may use the MS Excel function EXPONDIST($\frac{5}{60}$, 6.4, 1) to find $\Pr(Y \leq \frac{5}{60})$.

9. (a) $\mathbb{E}[X+2] = \mathbb{E}[X] + 2 = \frac{1}{3} + 2 = \frac{7}{3}$.

   (b) $\text{Var}(X+2) = \text{Var}(X) = (\frac{1}{3})^2 = \frac{1}{9}$.

   (c) $\Pr(X+2 > 1) = \Pr(X > -1) = 1$.

   (d) No, $X+2$ is not an exponential distribution. If it is, the square of its mean $\frac{7}{3}$ should be the variance. Nevertheless, $\left(\frac{7}{3}\right)^2 \neq \frac{1}{9}$ and thus we know $X+1$ is not an exponential distribution. Another way to understand this is to see that for the new random variable $X+2$, it is impossible for any value in $[0,2)$ to occur, which violates the pdf of any Exponential random variable.

10. Recall that in Problem 7 of Homework 5, we considered a random variable X which follows a binomial distribution with $n = 20$ and $p = 0.1$.

    (a) (3 points) If you use a normal random variable to approximate X, what should be the parameters of the normal random variable?
    In using a normal random variable $Y \sim \text{ND}(\mu, \sigma)$ to approximate the binomial random variable X, we should choose $\mu = np = 2$ and $\sigma = \sqrt{np(1-p)} \approx 1.34$.

    (b) We will use $\Pr(Y \leq 3)$ to approximate $\Pr(X \leq 3)$. We have

    $$\Pr(Y \leq 3) = \Pr\left(Z \leq \frac{3-2}{1.34}\right) = \Pr(Z \leq 0.7453) \approx 0.772.$$

    This approximation is bad because 0.772 is far from the exact probability 0.867.

    (c) We will use $\Pr(Y \leq 3.5)$ to approximate $\Pr(X \leq 3)$. We have

    $$\Pr(Y \leq 3.5) = \Pr\left(Z \leq \frac{3.5-2}{1.34}\right) = \Pr(Z \leq 1.118) \approx 0.868.$$

    This approximation is good because 0.868 is quite close to the exact probability 0.867.