Ishaan Taneja

Grading:
30 marks- Complete submission of Problem 2,5
70 marks- Problem 2

1. From the graph $G(10, \frac{x}{6})$ or from A that you constructed in the worksheet:
   (a) fill in the following table from the data in worksheet:

x	# Edges

   (b) Let E denote the number of edges in a realisation of $G(10, \frac{x}{6})$. Find the likelihood $L(x; E)$ that E edges occur in the random Graph $G(10, \frac{x}{6})$.
   (c) Find x^* that maximizes $L(x; E)$ with respect to x. You may assume $x \in [1, 5]$.
   (d) Substitute your value of E from Question 1, into the expression for x^*. Is the resulting x^* close to your chosen x ?

Solution: 1

(a) The data from worksheet:

x	# Edges
3	21

(b) $G(10, x/6)$ has 10 vertices ($n = 10$).
So the possible number of edges is ${}^{10}C_2 = 45$.
E= Number of edges in a realisation of $G(10, x/6)$
i.e. E is the event of occurrence of edges among the 45 possible edges.
Probability of success of each edge $= x/6$
$\therefore E \sim Binomial(45, x/6)$
The p.m.f of E is given by:

$$f_x(E) = \binom{45}{E} \left(\frac{x}{6}\right)^E \left(1 - \frac{x}{6}\right)^{45-E}.$$

(c) Since we are only interested in the argument that maximizes $f_x(E)$, we will first take log, and then maximize.

$$\log f_x(E) = \log\left[\binom{45}{E}\right] + E \log \frac{x}{6} + (45 - E) \log\left(1 - \frac{x}{6}\right)$$
$$\Rightarrow \frac{d \log f_x(E)}{d\,x} = E\frac{1}{\frac{x}{6}} \cdot \frac{1}{6} + (45 - E)\,\frac{1}{1 - \frac{x}{6}}\left(-\frac{1}{6}\right) \overset{set}{=} 0$$

$$
\begin{aligned}
&\Rightarrow E\frac{1}{\frac{x^*}{6}} = (45-E)\frac{1}{1-\frac{x^*}{6}}\\
&\Rightarrow \frac{1-\frac{x^*}{6}}{\frac{x}{6}} = \frac{45}{E} - 1\\
&\Rightarrow \frac{1}{\frac{x^*}{6}} = \frac{45}{E}\\
&\Rightarrow x^* = 6 \times \frac{E}{45}
\end{aligned}
$$

Also, $\dfrac{d^2 \log f_x(E)}{d\,x^2} < 0 \quad ;$ for $x = x^*$.

Thus, x^* is the point of maxima for $f_x(E)$.

(d) $E = 21$

$\Rightarrow x^* = 6 \times \dfrac{21}{45} = 2.8 \approx 3$

Hence, for our chosen $x(=3)$ is very close to 2.8.

2. Example 9.2.1.

Solution: 2

X_1,X_2,...,X_{10} are i.i.d Binomial(N,p)

Number of unknown parameters= 2

Empirical realisation $= 8, 7, 6, 11, 8, 5, 3, 7, 6, 9$

$$
\begin{aligned}
\mu_1(X_1, X_2, ..., X_{10}) &= \frac{1}{10}\sum_{i=1}^{10} X_i\\
&= \frac{1}{10}(8+7+6+11+8+5+3+7+6+9)\\
&= 7
\end{aligned}
$$

$$
\begin{aligned}
\mu_2(X_1, X_2, ..., X_{10}) &= \frac{1}{10}\sum_{i=1}^{10} X_i^2\\
&= \frac{1}{10}(8^2+7^2+6^2+11^2+8^2+5^2+3^2+7^2+6^2+9^2)\\
&= 53.4
\end{aligned}
$$

And, $m_1 = E(X) = Np$

$m_2 = E(X^2) = Np(1-p) + (Np)^2$

By method of moments:

$\mu_1(X_1, X_2, ..., X_{10}) = m_1 \Rightarrow Np = 7$ —(1)

And, $\mu_2(X_1, X_2, ..., X_{10}) = m_2 \Rightarrow Np(1-p) + (Np)^2 = 53.4$ —(2)

Substituting (1) in (2); we get:

$$
\begin{aligned}
Np(1-p) + (Np)^2 &= 53.4\\
7(1-p) + 49 &= 53.4
\end{aligned}
$$

$$7 - 7p = 4.4$$
$$\therefore p = 0.371$$

Putting value of p in (1); we get:
$N = \frac{7}{0.371} = 18.87 \approx 19$

Thus, the method of moments estimates that the distribution from which the sample came is $Binomial(19, 0.371)$.

3. Example 9.2.2.
   Solution: 3
   X_1,X_2,...,X_n are i.i.d Normal(μ, σ)
   $\mu_1(X_1, X_2, ..., X_n) = \frac{1}{n}\sum_{i=1}^{n} X_i$
   $\mu_2(X_1, X_2, ..., X_n) = \frac{1}{n}\sum_{i=1}^{n} X_i^2$
   And, $m_1 = E(X) = \mu$
   $m_2 = E(X^2) = Var(X) + (E(X))^2 = \sigma^2 + \mu^2$
   By method of moments:
   $\mu_1(X_1, X_2, ..., X_n) = m_1 \Rightarrow \mu = \frac{1}{n}\sum_{i=1}^{n} X_i$ —(1)
   And, $\mu_2(X_1, X_2, ..., X_n) = m_2 \Rightarrow \sigma^2 + \mu^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$ —(2)
   Substituting (1) in (2); we get:
$$\sigma^2 + \mu^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$
$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2$$
$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$
$$\sigma^2 = \frac{n-1}{n}S^2 \quad ; \text{where, } S^2 = \text{Sample variance}$$
$$\therefore \sigma = \sqrt{\frac{n-1}{n}}S$$
   Thus, the method of moment estimators for μ and σ is $\overline{X}$ and $\sqrt{\frac{n-1}{n}}S$ respectively.

4. Example 9.3.2.
   Solution: 4
   X_1,X_2,...,X_n are i.i.d Normal$(p, 1)$
   $f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{(x-p)^2}{2}}$
   The likelihood function is given by:

$$
\begin{aligned}
L(p; X_1, ..., X_n) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{(X_i - p)^2}{2}} \\
&= \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{\sum_{i=1}^{n}(X_i - p)^2}{2}}
\end{aligned}
$$

To find MLE, treating the given realisation X_1,X_2,...,X_n as fixed, we needs to maximise L as a function of p.
This is equivalent to finding the minimum of $g : R \to R$ given by

$$
g(p) = \sum_{i=1}^{n}(X_i - p)^2
$$

Now,

$$
\begin{aligned}
g'(p) &= -2\sum_{i=1}^{n}(X_i - p) \overset{set}{=} 0 \\
n\hat{p} &= \sum_{i=1}^{n} X_i \\
\therefore \hat{p} &= \overline{X}
\end{aligned}
$$

Also, $g"(p) > 0$ for $p = \hat{p}$
Hence, the MLE of p is given by $\hat{p} = \overline{X}$

5. Example 9.3.3.
Solution: 5
X_1,X_2,...,X_n are i.i.d Bernoulli(p)

$$
f(x|p) = \begin{cases} p^x(1-p)^{1-x} & ; \text{if } x \in 0, 1 \\ 0 & ; \text{otherwise} \end{cases}
$$

The likelihood function is given by

$$
\begin{aligned}
L(p; X_1, ..., X_n) &= \prod_{i=1}^{n} p^{X_i}(1-p)^{1-X_i} \\
&= p^{\sum_{i=1}^{n} X_i}(1-p)^{n-\sum_{i=1}^{n} X_i}
\end{aligned}
$$

To find the MLE, treating the given realisation as fixed, we need to maximize L as a function of p.
The log likelihood function is:

$$
T(p; X_1, ..., X_n) = lnL(p; X_1, ..., X_n)
$$

$$
T(p; X_1, ..., X_n) = \begin{cases} \ln\left(\frac{p}{1-p}\right)a + n\ln(1-p) & ; \text{if } \sum_{i=1}^{n} X_i = a,\ 0 < a < n \\ n\ln(1-p) & ; \text{if } \sum_{i=1}^{n} X_i = 0 \\ n\ln p & ; \text{if } \sum_{i=1}^{n} X_i = n \end{cases}
$$

Case 1: If $0 < \sum_{i=1}^{n} X_i < n$

$$\frac{d\,T(p; X_1, ..., X_n)}{d\,p} = \frac{1}{p(1-p)} \sum_{i=1}^{n} X_i - \frac{n}{1-p} \overset{set}{=} 0$$

$$\frac{1}{\hat{p}(1-\hat{p})} \sum_{i=1}^{n} X_i = \frac{n}{1-\hat{p}}$$
$$\frac{1}{\hat{p}} \sum_{i=1}^{n} X_i = n$$
$$\therefore \hat{p} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Case 2: If $\sum_{i=1}^{n} X_i = 0$
In this case, T is a decreasing function of p and maximum occurs at $p = 0$.

Case 3: If $\sum_{i=1}^{n} X_i = n$
In this case, T is an increasing function of p and maximum occurs at $p = 1$.

Therefore, we can conclude that $\hat{p} = \dfrac{\sum_{i=1}^{n} X_i}{n}$.

6. Example 9.4.2.
   Solution: 6
   X_1,X_2,...,X_{16} are i.i.d X
   Sample mean ; $\overline{X} = 10.2$
   Sample standard deviation ; $S = 3.0$
   In this case, the confidence level ;$\beta = 0.95$.
   So we must find the value of a for:

$$P(|\overline{X} - \mu| < a) = 0.95$$
$$P\left(\frac{|\overline{X} - \mu|}{S/\sqrt{n}} < \frac{a}{S/\sqrt{n}}\right) = 0.95$$
$$P\left(|T| < \frac{4a}{3}\right) = 0.95$$
$$P\left(-\frac{4a}{3} < T < \frac{4a}{3}\right) = 0.95 \quad \text{—(1)}$$

   Also, we know that for t-distribution:

$$P(-2.13 < T < 2.13) = 0.95 \quad \text{—(2)}$$

   On comparing (1) and (2), we get:

$$\frac{4a}{3} = 2.13$$
$$\therefore a \approx 1.60$$

   Hence, 95% confidence interval for the actual mean of the distribution X is $(8.6, 11.8)$.