STAT 330
Tutorial 9

Mengqi (Molly) Cen
Department of Statistics & Actuarial Science

Nov 16th/ Nov 18th

Outline

1. **Convergence in Probability**

2. **Convergence in Distribution**

3. **Maximum Likelihood Methods**

Convergence in Probability

5.1.3. Let W_n denote a random variable with mean μ and variance b/n^p, where $p > 0$, μ, and b are constants (not functions of n). Prove that W_n converges in probability to μ.
Hint: Use Chebyshev's inequality.

Review:

We say a sequence of random variables (rvs) $\{Y_n : n = 1, 2, \ldots\}$ converges in probability to rv Y if, for $\forall \varepsilon > 0$,

$$\lim_{n\to\infty} P(|Y_n - Y| \geq \varepsilon) = 0.$$

Denoted it by "$Y_n \to Y$ in probability" as $n\to\infty$, or "$Y_n \xrightarrow{P} Y$ as $n\to\infty$".

Answer:

By Chebyshev's inequality

$$P(|W_n - \mu| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} Var[W_n] = \frac{b}{\varepsilon^2 n^p}, \text{ for } \varepsilon > 0.$$

Since $p > 0$,

$$\lim_{n\to\infty} \frac{b}{\varepsilon^2 n^p} = 0.$$

Therefore, $\lim_{n\to\infty} P(|W_n - \mu| \geq \varepsilon) = 0$, which shows that $W_n \xrightarrow{P} \mu$.

Convergence in Probability

5.1.5. Let $X_1, \ldots, X_n$ be iid random variables with common pdf

$$f(x) = \begin{cases} e^{-(x-\theta)} & x > \theta \quad -\infty < \theta < \infty \\ 0 & \text{elsewhere.} \end{cases} \tag{5.1.3}$$

This pdf is called the **shifted exponential**. Let $Y_n = \min\{X_1, \ldots, X_n\}$. Prove that $Y_n \to \theta$ in probability, by first obtaining the cdf of Y_n.

<u>Answer</u>:

For any $a > \theta$,

$P(X \geq a) = \int_a^{\infty} e^{-(x-\theta)} dx = e^{-(a-\theta)}$.

Since $Y_n = \min\{X_1, \ldots, X_n\}$,

$$\begin{aligned} P(Y_n \geq a) &= P(X_1 \geq a, \ldots, X_n \geq a) \\ &= P(X_1 \geq a) \cdots P(X_n \geq a) \\ &= [e^{-(a-\theta)}]^n \\ &= e^{-n(a-\theta)}. \end{aligned}$$

Since $Y_n > \theta$,

$$\begin{aligned} P(|Y_n - \theta| \geq \varepsilon) &= P(Y_n - \theta \geq \varepsilon) \\ &= P(Y_n \geq \varepsilon + \theta) \\ &= e^{-n(\varepsilon + \theta - \theta)} \\ &= e^{-n\varepsilon}, \text{ for } \varepsilon > 0. \end{aligned}$$

Therefore,

$\lim_{n \to \infty} P(|Y_n - \theta| \geq \varepsilon) = \lim_{n \to \infty} e^{-n\varepsilon} = 0$,

which shows that $Y_n \xrightarrow{P} \theta$.

Convergence in Distribution

5.2.2. Let Y_1 denote the minimum of a random sample of size n from a distribution that has pdf $f(x) = e^{-(x-\theta)}$, $\theta < x < \infty$, zero elsewhere. Let $Z_n = n(Y_1 - \theta)$. Investigate the limiting distribution of Z_n.

Review:

Consider rv $X \sim F_X(\cdot)$. We call a sequence of rvs $\{X_n\}$ converges in distribution to X if

$$\lim_{n\to\infty} F_{X_n}(x) = \lim_{n\to\infty} P(X_n \le x) = F(x)$$

for all $x \in C(F_X)$, the set of all continuous points of $F_X(\cdot)$. Denote it by $X_n \xrightarrow{D} X$.

Answer:

Previous question has showed that

$$P(Y_1 \geq y) = e^{-n(y-\theta)} \text{ for } y > \theta.$$

Since $\theta < x < \infty$, $0 < z < \infty$.

Thus for $t \le 0$, $F_{Z_n}(t) = P(Z_n \le t) = 0$.

For $t > 0$,
$$\begin{aligned} F_{Z_n}(t) &= P(Z_n \le t) \\ &= P(n(Y_1 - \theta) \le t) \\ &= P\left(Y_1 \le \frac{t}{n} + \theta\right) \\ &= 1 - P\left(Y_1 > \frac{t}{n} + \theta\right) \\ &= 1 - e^{-t}. \end{aligned}$$

We see that F_{Z_n} is the cdf for the exponential distribution with mean $\mu = 1$. Therefore, $\{Z_n\}$ converges in distribution to an exponential distribution with mean $\mu = 1$.

Convergence in Distribution

5.2.7. Let X_n have a gamma distribution with parameter $\alpha = n$ and β, where β is not a function of n. Let $Y_n = X_n/n$. Find the limiting distribution of Y_n.

Review:

Consider the sequence of rvs $\{X_n\}$ with the mgf of X_n as $M_{X_n}(t)$ for $-h<t<h$, and a rv X with mgf $M_X(t)$ for $-h<t<h$.

If $\lim_{n\to\infty} M_{X_n}(t) = M_X(t)$ for $-h<t<h$,

then $X_n \xrightarrow{D} X$.

Answer:

$$M_{X_n}(t) = E(e^{tX_n})$$
$$= \frac{1}{\Gamma(n)\beta^n}\int_0^\infty e^{tx} x^{n-1} e^{-\frac{x}{\beta}}\,dx$$
$$= \frac{1}{\Gamma(n)\beta^n}\int_0^\infty x^{n-1} e^{-\frac{x}{\frac{\beta}{1-\beta t}}}\,dx$$
$$= \frac{1}{\Gamma(n)\beta^n}\Gamma(n)\left(\frac{\beta}{1-\beta t}\right)^n$$
$$= \left(\frac{1}{1-\beta t}\right)^n.$$

$$M_{Y_n}(t) = E(e^{tY_n}) = E(e^{\frac{t}{n}X_n}) = M_{X_n}\left(\tfrac{t}{n}\right)$$
$$= \left(\frac{1}{1-\beta t/n}\right)^n.$$

$$\lim_{n\to\infty} M_{Y_n}(t) = \frac{1}{\lim_{n\to\infty}(1-\beta t/n)^n}$$
$$= e^{\beta t},$$

which is the mgf of constant β.

Convergence in Distribution

5.3.8. Let Y be $b(n, 0.55)$. Find the smallest value of n which is such that (approximately) $P(Y/n > \frac{1}{2}) \geq 0.95$.

Review:

Central Limit Theorem

If $X_1, \dots, X_n$ are iid with mean μ and variance σ^2,

$$Y_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right) \xrightarrow{D} N(0,1),$$

as $n \to \infty$.

That is, $\dfrac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} \xrightarrow{D} N(0,1)$, as $n \to \infty$.

Answer:

Given $Y \sim b(n, 0.55)$,

we have $E(Y) = 0.55n$,

$Var(Y) = n \cdot 0.55 \cdot 0.45 = 0.2475n$

$$P\left(\frac{Y}{n} > \frac{1}{2}\right)$$
$$= 1 - P\left(\frac{Y}{n} \leq \frac{1}{2}\right)$$
$$= 1 - P\left(\frac{Y - 0.55n}{\sqrt{0.2475n}} \leq \frac{0.5n + 0.5 - 0.55n}{\sqrt{0.2475n}}\right)$$
$$\approx 1 - P\left(Z \leq \frac{0.5 - 0.05n}{\sqrt{0.2475n}}\right) \geq 0.95$$

$$\therefore \Phi\left(\frac{0.5 - 0.05n}{\sqrt{0.2475n}}\right) \leq 0.5$$

Thus, $\dfrac{0.5 - 0.05n}{\sqrt{0.2475n}} \leq -1.645$.

Solve for n we get $n = 290$.

Maximum Likelihood Estimation

6.1.6. Let the table

x	0	1	2	3	4	5
Frequency	6	10	14	13	6	1

represent a summary of a sample of size 50 from a binomial distribution having $n = 5$. Find the mle of $P(X \geq 3)$.

Review:

The MLE $\hat{\theta}$ is the value of the population parameter θ that maximizes the likelihood function:

$$L(\hat{\theta}|\text{data}) = \max_{\theta \in \Omega} L(\theta|\text{data}).$$

Answer:

Let $X_1, X_2, \ldots, X_n$ are n iid variables with binomial distribution with parameters n and p. Then we have that $P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$.

Then the likelihood function is: $L(p|X_1,\ldots,X_m) = \prod_{i=1}^{m} \binom{n}{x_i} p^{x_i} (1-p)^{n-x_i}$ for $i=1,\ldots,m$.

$l(\hat{p}|X_1,\ldots,X_m) = \sum_{i=1}^{m} \ln\binom{n}{x_i} + \sum_{i=1}^{m} x_i \ln(p) + \sum_{i=1}^{m}(n-x_i)\ln(1-p) = \sum_{i=1}^{m} \ln\binom{n}{x_i} + \sum_{i=1}^{m} x_i \ln(p) + (mn - \sum_{i=1}^{m} x_i)\ln(1-p)$.

Using derivative of function equals 0 to find MLE:

$$\frac{\partial l}{\partial p} = \frac{\sum_{i=1}^{m} x_i}{p} - \frac{mn - \sum_{i=1}^{n} x_i}{1-p} = 0,$$

We get $\hat{p} = \frac{1}{n}\bar{x}$.

The second derivative of function is $-\frac{\sum_{i=1}^{m} x_i}{p^2} - \frac{mn - \sum_{i=1}^{m} x_i}{(1-p)^2}$.

Since $n \geq x_i$, for $i = 1, \ldots, m$, $\frac{\partial^2 l}{\partial p} \geq 0$ so $\hat{p} = \frac{1}{n}\bar{x}$ is the MLE.

$$\hat{p} = \frac{\bar{x}}{n} = \frac{0\cdot 6 + 1\cdot 10 + 2\cdot 14 + 3\cdot 13 + 4\cdot 6 + 5\cdot 1}{50\cdot 5} = 0.424.$$

$$\theta = P(X \geq 3) = P(X=3) + P(X=4) + P(X=5)$$

$$\hat{\theta} = \binom{5}{3} 0.424^3 \cdot 0.576^2 + \binom{5}{4} 0.424^4 \cdot 0.576^1 + \binom{5}{5} 0.424^5 \cdot 0.576^0$$

$$= 0.36.$$

Maximum Likelihood Estimation

6.2.8. Let X be $N(0,\theta)$, $0<\theta<\infty$.

(a) Find the Fisher information $I(\theta)$.

(b) If $X_1, X_2, \ldots, X_n$ is a random sample from this distribution, show that the mle of θ is an efficient estimator of θ.

Review:

The Fisher Information is $FI(\theta) = E\left[\left(\frac{\partial \log f(X;\theta)}{\partial\theta}\right)^2\right]$, provided the expectation exists.

Note that $FI(\theta) = Var\left(\frac{\partial \log f(X;\theta)}{\partial\theta}\right) = -E\left[\frac{\partial^2 \log f(X;\theta)}{\partial^2\theta}\right]$.

Answer:

a) Given $X \sim N(0,\theta)$, $0<\theta<\infty$, $f(x;\theta) = \frac{1}{\sqrt{2\pi\theta}} \exp\{-\frac{1}{2}\frac{x^2}{\theta}\}$

$\log f(x;\theta) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\theta) - \frac{x^2}{2\theta}$.

$\frac{\partial \log f(x;\theta)}{\partial\theta} = -\frac{1}{2\theta} + \frac{x^2}{2\theta^2}$.

$\frac{\partial^2 \log f(x;\theta)}{\partial^2\theta} = \frac{1}{2\theta^2} - \frac{x^2}{\theta^3}$.

$FI(\theta) = -E\left[\frac{\partial^2 \log f(X;\theta)}{\partial^2\theta}\right] = -E\left[\frac{1}{2\theta^2} - \frac{x^2}{\theta^3}\right] = -\frac{1}{2\theta^2} + \frac{1}{\theta^2}E\left[\frac{x^2}{\theta}\right]$

Since $\frac{X^2}{\theta} \sim \chi^2_1$, $E\left[\frac{X^2}{\theta}\right] = 1$, $FI(\theta) = \frac{1}{2\theta^2}$.

Review:

An unbiased estimator Y with a random sample of size n is called efficient if $Var(Y) = \frac{1}{nFI(\theta)}$.

Answer:

b) $l(\theta) = \sum_{i=1}^{n} \log f(x_i;\theta) = -\frac{1}{2}\sum_{i=1}^{n}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}\log(\theta) - \sum_{i=1}^{n}\frac{x_i^2}{2\theta}$

$= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\theta) - \frac{1}{2\theta}\sum_{i=1}^{n}x_i^2$.

Using derivative of function equals 0 to find MLE:

$\frac{\partial l}{\partial\theta} = -\frac{n}{2\theta} + \frac{\sum_{i=1}^{n}x_i^2}{2\theta^2} = 0$, we get $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}x_i^2$.

$\frac{\partial^2 l}{\partial\theta^2} = \frac{n}{2\theta^2} - \frac{\sum_{i=1}^{n}x_i^2}{\theta^3}$

After pluging $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}x_i^2$, we get $\frac{\partial^2 l}{\partial\theta^2} = -\frac{n^3}{2(\sum_{i=1}^{n}x_i^2)} < 0$

Thus, $\hat{\theta}$ is the MLE.

Since $\frac{\sum_{i=1}^{n}X_i^2}{\theta} \sim \chi^2_n$,

$Var(\hat{\theta}) = \frac{\theta^2}{n^2} Var\left(\frac{\sum_{i=1}^{n}X_i^2}{\theta}\right) = \frac{2\theta^2}{n} = \frac{1}{nFI(\theta)}$.

Questions