

MATH324 (Statistics) – Lecture Notes

McGill University

Masoud Asgharian

Winter 2019

Contents

1	Lecture 1	3
2	Lecture 2	3
3	Lecture 3	4
3.1	Markov's Inequality	4
3.2	Tchebyshev's Inequality	4
3.3	Application to Voting	7
4	Lecture 4	9
5	Lecture 5	9
6	Lecture 6	9
7	Lecture 7	9
8	Lecture 8	9
9	Lecture 9	9
10	Lecture 10	9

11 Lecture 11	9
12 Lecture 12	9

1 Lecture 1

2 Lecture 2

3 Lecture 3

3.1 Markov's Inequality

Let X be a random variable and h be a **non-negative** function; ie:

$$h : R \rightarrow R^+ \cup \{0\} = [0, \infty)$$

Suppose $E(h(X)) < \infty$, then for some $\lambda > 0$, we have:

$$P(h(X) \geq \lambda) \leq \frac{E[h(X)]}{\lambda} \quad (1)$$

Proof. Suppose X is a continuous random variable:

$$\begin{aligned} E[h(x)] &= \int_x h(x) f_X(x) dx \\ &= \left(\int_{x:h(x) \geq \lambda} h(x) f_X(x) dx + \int_{x:h(x) < \lambda} h(x) f_X(x) dx \right) \\ &\geq \int_{x:h(x) \geq \lambda} h(x) f_X(x) dx \quad \text{since } h \geq 0 \\ &\geq \lambda \int_{x:h(x) \geq \lambda} f_X(x) dx = \lambda P(h(X) \geq \lambda) \\ \Rightarrow P(h(X) \geq \lambda) &\leq \frac{E(h(X))}{\lambda} \end{aligned}$$

The proof for the discrete case is similar. □

3.2 Tchebyshev's Inequality

Tchebyshev's Inequality is a special case of Markov's Inequality. Consider $h(x) = (x - \mu)^2$, then:

$$\begin{aligned} P(|X - \mu| \geq \lambda) &= P((X - \mu)^2 \geq \lambda^2) \\ &\leq \frac{E[(X - \mu)^2]}{\lambda^2} \quad \text{if } E[(X - \mu)^2] < \infty \end{aligned}$$

Let $\mu = E(X)$, then $E[(X - \mu)^2] = \text{Var}(X)$ denoted by σ_x^2 . We therefore have:

$$P(|X - \mu_x| \geq \lambda) \leq \frac{\sigma_x^2}{\lambda^2} \quad \text{where } \mu_x = E(X) \quad (2)$$

Now consider $\lambda = K\sigma_x$ where K is a known number. Then:

$$P(|X - \mu_x| \geq K\sigma_x) \geq \frac{\sigma_x^2}{K^2\sigma_x^2} = \frac{1}{K^2} \quad (3)$$

This is called **Tchbyshev's Inequality**.

Example 3.1. Suppose $K = 3$.

$$P(|X - \mu_x| \geq 3\sigma_x) \leq \frac{1}{9}$$

In other words, at least 88% of the observations are within 3 standard deviation from the population mean.

Going back to the our example:

$$X_i \sim (\mu, 1) \quad , \quad \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

We want to study $P(\epsilon \geq \delta) = P(|\bar{X}_n - \mu| \geq \delta)$, first we note that:

$$E(X_i) = \mu \quad , \quad i = 1, 2, \dots, n$$

Then:

$$\begin{aligned} E(\bar{X}_n) &= E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) \\ &= \frac{1}{n} \sum_{i=1}^n \mu = \frac{1}{n} (n\mu) = \frac{1}{n} \cdot (n\mu) \\ &= \mu \end{aligned} \quad (*)$$

Thus, using (2) we have:

$$P(|\bar{X}_n - \mu| \geq \delta) \leq \frac{Var(\bar{X}_n)}{\delta^2}$$

Now:

$$\begin{aligned} Var(\bar{X}_n) &= Var\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} Var\left(\sum_{i=1}^n X_i\right) \\ &= \frac{1}{n^2} \left[\sum_{i=1}^n Var(X_i) + \sum_{1 \leq i < j \leq n} \sum_{1 \leq i < j \leq n} Cov(X_i, X_j) \right] \quad \text{using Thm 5.12(b) - page 271} \\ &= \frac{1}{n^2} \sum_{i=1}^n Var(X_i) \quad \text{since } \prod_{i=1}^n X_i \\ &= \frac{1}{n^2} n Var(X) = \frac{Var(X)}{n} \quad \text{since } x_i \text{ s are identically distributed} \\ &= \frac{\delta_X^2}{n} \end{aligned} \quad (**)$$

In our case $X \sim N(\mu, 1)$ so $Var(X) = \delta_X^2 = 1$. Thus $Var(\bar{X}_n) = \frac{1}{n}$

Remark. $X \coprod Y \implies Cov(X, Y) = 0$. Note that:

$$X \coprod Y \implies E[g_1(X)g_2(Y)] = E[g_1(X)].E[g_2(Y)]$$

in particular:

$$X \coprod Y \implies E[XY] = E[X].E[Y]$$

on the other hand:

$$Cov(X, Y) = E[XY] - E(X)E(Y)$$

thus:

$$X \coprod Y \implies Cov(X, Y) = 0.$$

recall that $X \coprod Y$ means X and Y are independent, i.e. $f_{X,Y}(x, y) = f_X(x)f_Y(y)$
where $f_{X,Y}$, f_X and f_Y represent respectively the

We therefore have:

$$P(|\bar{X}_n - \mu| \geq \delta) \leq \frac{1}{n\delta^2} \quad (4)$$

Using (4) and the sample size, n , we can find an upper bound for the proportion of deviations which are greater than a given threshold δ .

We can also use (4) for Sample Size Deterministic:

Suppose δ is given and we want $P(|\bar{X}_n - \mu| \geq \delta) \leq \beta$ where β is also given.

Then setting $\frac{1}{n\delta^2} = \beta$, we can estimate $n \approx \frac{1}{\beta\delta^2}$.

3.3 Application to Voting

Define $X_i = \begin{cases} 1 & \text{NDP} \\ 0 & \text{otherwise} \end{cases}$. Associated to each eligible voter in Canada we

have a binary variable X . Let $p = P(X = 1)$. So p represents the proportion of eligible voters who favor *NDP*. Of interest is often estimation of p . Suppose we have a sample of size n , X_1, X_2, \dots, X_n .

$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ is the sample proportion; The counterpart of p which can be denoted by \hat{p} . Note that:

$$\mu_X = E(X) = 1 \times P(X = 1) + 0 \times P(X = 0) = 1 - p + 0 \times (1 - p) = p$$

and:

$$E(X^2) = 1^2 \times P(X = 1) + 0^2 \times P(X = 0) = 1 - p + 0 \times (1 - p) = p$$

From (*) and (**) we find that :

$$E(\hat{p}_n) = E(\bar{X}_n) = \mu_X = p$$

and:

$$Var(\hat{p}_n) = E(\bar{X}_n) = \frac{Var(X)}{n} = \frac{\sigma_X^2}{n} = \frac{p(1-p)}{n}$$

Thus using (2), we have:

$$P(|\hat{p}_n - p| \geq \delta) \leq \frac{Var(\hat{p}_n)}{\delta^2} = \frac{p(1-p)}{n\delta^2}$$

Note that the above bound on the probability of derivation depends on p which is *unknown*. We however notice that $p(1-p) \leq \frac{1}{4}$.

Define $\zeta(x) = x(1-x)$ for $0 < x < 1$. Then:

$$\begin{aligned} \zeta'(x) = 1 - 2x &\implies \zeta'(x) = 0 \implies x = \frac{1}{2} \\ \zeta''(\frac{1}{2}) = -2 &\implies x = \frac{1}{2} \quad \text{which is a **maximizer**} \\ \zeta(\frac{1}{2}) = \frac{1}{2}(1 - \frac{1}{2}) &= \frac{1}{4} \end{aligned}$$

(Note that $\zeta''(x) = -2$ for all $0 < x < 1$)

We therefore find:

$$P(|\hat{p}_n - p| \geq \delta) \leq \frac{1}{4n\delta^2} \tag{5}$$

Using (5) and a given sample size n we can find an upper bound for the probability of derivation by δ and the amount for any given δ .

We can also use (5) for sample size deterministic for a size bound β and derivative δ as follows:

$$\frac{1}{4n\delta^2} = \beta \implies n \geq \frac{1}{4\beta\delta^2}$$

This is of course conservative since $p(1-p) \leq \frac{1}{4}$.

4 Lecture 4

5 Lecture 5

6 Lecture 6

7 Lecture 7

8 Lecture 8

9 Lecture 9

10 Lecture 10

11 Lecture 11

12 Lecture 12