

## 36-705 Intermediate Statistics.

### Homework 3.

Date: 18th June 2020.

Correction status: pending.

1) Let  $\mathcal{C} = \mathcal{A} \cup \mathcal{B}$ . Show that

$$s_n(\mathcal{C}) \leq s_n(\mathcal{A}) + s_n(\mathcal{B})$$

where  $s_n$  denotes the shattering number.

2) Let  $\mathcal{C} = \{A \cup B; A \in \mathcal{A}, B \in \mathcal{B}\}$ . Show that:

$$s_n(\mathcal{C}) \leq s_n(\mathcal{A})s_n(\mathcal{B})$$

3) Chapter 5, problem 2.

Let  $X_1, X_2, \dots$  be a sequence of random variables. Show that  $X_n \xrightarrow{\text{qm}} b$  if and only if:

$$\lim_{n \rightarrow \infty} \mathbb{E}[X_n] = b \quad \text{and} \quad \lim_{n \rightarrow \infty} \text{Var}[X_n] = 0$$

Denoting proposition (I) as:

$$X_n \xrightarrow{\text{qm}} b$$

And proposition (II) as:

$$\lim_{n \rightarrow \infty} \mathbb{E}[X_n] = b \quad \text{and} \quad \lim_{n \rightarrow \infty} \text{Var}[X_n] = 0$$

To show that (I)  $\implies$  (II):

In order for convergence in quadratic mean,  $X_n \xrightarrow{\text{qm}} b$ , by definition, we have that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \mathbb{E}[(X_n - b)^2] = 0 \\ \implies & \lim_{n \rightarrow \infty} \mathbb{E}[X_n^2 - 2bX_n + b^2] = 0 \\ \implies & \lim_{n \rightarrow \infty} \mathbb{E}[X_n^2] - 2b\mathbb{E}[X_n] + b^2 = 0 \\ \implies & \lim_{n \rightarrow \infty} \text{Var}[X_n] + \mathbb{E}[X_n]^2 - 2b\mathbb{E}[X_n] + b^2 = 0 \\ \implies & \lim_{n \rightarrow \infty} \text{Var}[X_n] + \lim_{n \rightarrow \infty} (\mathbb{E}[X_n] - b)^2 = 0 \end{aligned}$$

Where we have substituted the 2nd moment for the sum of the variance and squared mean to get from the 3rd to the 4th equality.

A property of both the variance  $\text{Var}[X_n]$  and the term  $(\mathbb{E}[X_n] - b)^2$  is that they are both non-negative. And because the RHS is 0, in order for equality to hold, we must have that both limits on the LHS be equal to 0, which occurs when

$$\lim_{n \rightarrow \infty} \mathbb{E}[X_n] = b \quad \text{and} \quad \lim_{n \rightarrow \infty} \text{Var}[X_n] = 0$$

Showing (II)  $\implies$  (I):

Because it is the case that

$$\lim_{n \rightarrow \infty} \mathbb{E}[X_n] = b \quad \text{and} \quad \lim_{n \rightarrow \infty} \text{Var}[X_n] = 0$$

And also because  $\text{Var}[X_n] = \mathbb{E}[X_n^2] - \mathbb{E}[X_n]^2$ , we have the following limit on the 2nd moment

$$\lim_{n \rightarrow \infty} \mathbb{E}[X_n^2] = b^2$$

We now consider

$$\lim_{n \rightarrow \infty} \mathbb{E}[(X_n - b)^2] = \lim_{n \rightarrow \infty} \mathbb{E}[X_n^2] - 2b \lim_{n \rightarrow \infty} \mathbb{E}[X_n] + b^2 = b^2 - 2b(b) + b^2 = 0$$

Which is the required result.

4) Chapter 5, problem 5.

Let  $X_1, \dots, X_n \sim \text{Bernoulli}(p)$ . Prove that:

$$\frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow{P} p \quad \text{and} \quad \frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow{\text{qm}} p$$

Bernoulli random variables have mean  $\mathbb{E}[X_i] = p$  and variance  $\text{Var}[X_i] = p(1 - p)$ .

We define a new random variable  $Y_i = X_i^2$ , and instead consider  $Y_1, Y_2, \dots, Y_n$ .

The mean of  $Y_i$  is given by

$$\mathbb{E}[Y_i] = \mathbb{E}[X_i^2] = \text{Var}[X_i] + \mathbb{E}[X_i]^2 = p(1 - p) + p^2 = p$$

As each of the  $X_i$  are Bernoulli random variables taking values of either 0 or 1, the random variables  $Y_i = X_i^2$  are bounded within the interval  $[0, 1]$ .

Applying Hoeffding's inequality to the sequence of sample means  $\bar{Y}_n$  and considering when  $n \rightarrow \infty$  we have that  $\forall \epsilon > 0$ ,

$$P(|\bar{Y}_n - p| \geq \epsilon) \leq 2e^{-2n\epsilon^2} \rightarrow 0$$

As  $Y_i = X_i^2$ , we have that  $\bar{Y}_n = \frac{1}{n} \sum_{i=1}^n X_i^2$ , and hence we have the required result concerning convergence in probability of the 2nd sample moment:

$$\bar{Y}_n \xrightarrow{P} p \implies \frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow{P} p$$

To show convergence in quadratic mean, we consider the expression  $\mathbb{E}[(\bar{Y}_n - p)^2]$ :

$$\mathbb{E}[(\bar{Y}_n - p)^2] = \mathbb{E}[(\bar{Y}_n^2 - 2p\bar{Y}_n + p^2)] = \mathbb{E}[\bar{Y}_n^2] - 2p\mathbb{E}[\bar{Y}_n] + p^2$$

Substituting the variance and mean of the sample mean in place of its 2nd moment we have that

$$\mathbb{E}[(\bar{Y}_n - p)^2] = \text{Var}[\bar{Y}_n] + \mathbb{E}[\bar{Y}_n]^2 - 2p\mathbb{E}[\bar{Y}_n] + p^2$$

As  $\bar{Y}_n$  is a sample mean, we have that  $\text{Var}[\bar{Y}_n] = \text{Var}[Y_i]/n$ , and evaluating the expression for the variance of  $Y_i$ , we have that

$$\text{Var}[Y_i] = \mathbb{E}[Y_i^2] - \mathbb{E}[Y_i]^2 = \mathbb{E}[X_i^4] - p^2$$

We now need to evaluate the fourth moment of the Bernoulli random variable  $X_i$ . We find the moment generating function  $\psi_{X_i}(t)$  as follows:

$$\begin{aligned} \psi_{X_i}(t) &= \mathbb{E}[e^{tX}] = \int e^{tX} dF(x) = \sum_{x_i} e^{tx_i} P(X_i = x_i) = e^t P(X_i = 1) + e^0 P(X_i = 0) \\ &\implies \psi_{X_i}(t) = pe^t + (1 - p) \end{aligned}$$

As  $\frac{d}{dt}e^t = e^t$  for all  $t$ , we have that the  $k$ th derivative of the moment generating function has the form

$$\psi_{X_i}^{(k)}(t) = pe^t \quad \forall k \geq 1$$

Evaluating at  $t = 0$ , we find that for Bernoulli random variables, all  $k$ th moments have the form  $\psi_{X_i}^{(k)}(0) = \mathbb{E}[X_i^k] = p$  for  $k \geq 1$ .

We then have that

$$\text{Var}[\bar{Y}_n] = \frac{\text{Var}[Y_i]}{n} = \frac{p - p^2}{n} = \frac{p(1 - p)}{n}$$

As  $\mathbb{E}[\bar{Y}_n] = \mathbb{E}[Y_i] = p$ , we have that

$$\begin{aligned} \mathbb{E}[(\bar{Y}_n - p)^2] &= \text{Var}[\bar{Y}_n] + \mathbb{E}[\bar{Y}_n]^2 - 2p\mathbb{E}[\bar{Y}_n] + p^2 = \frac{p(1 - p)}{n} + p^2 - 2p(p) + p^2 \\ &= \frac{p(1 - p)}{n} \end{aligned}$$

As  $n \rightarrow \infty$ , we have that  $\mathbb{E}[(\bar{Y}_n - p)^2] = \frac{p(1-p)}{n} \rightarrow 0$ , which is the required result.

*It was only observed after completing the problem that for Bernoulli random variables,  $X_i^2 = X_i$ , thereby rendering the need for computations using moment generating functions redundant in the calculation of  $\text{Var}[Y_i] = \text{Var}[X_i^2]$ , as it is the case that  $\mathbb{E}[X_i^2] = \mathbb{E}[X_i] = p$ .*

5) Chapter 5, problem 12.

Let  $X_1, X_2, \dots$  be random variables that are positive and integer valued.

Show that  $X_n \xrightarrow{D} X$  if and only if:

$$\lim_{n \rightarrow \infty} P(X_n = k) = P(X = k)$$

6) Chapter 5, problem 15.

Let

$$\begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix}, \begin{pmatrix} X_{12} \\ X_{22} \end{pmatrix}, \dots, \begin{pmatrix} X_{1n} \\ X_{2n} \end{pmatrix}$$

be IID random vectors with mean  $\boldsymbol{\mu} = (\mu_1, \mu_2)$  and variance  $\Sigma$ . Assume that  $\mu_2 \neq 0$ . Then let

$$\bar{X}_1 = \frac{1}{n} \sum_{i=1}^n X_{1i}, \quad \bar{X}_2 = \frac{1}{n} \sum_{i=1}^n X_{2i}$$

and define  $Y_n = \bar{X}_1 / \bar{X}_2$ . Find the limiting distribution of  $Y_n$ .

Defining the sample mean vector  $\bar{\mathbf{X}}_n \in \mathbb{R}^2$  as follows:

$$\bar{\mathbf{X}}_n = \begin{pmatrix} \bar{X}_1 \\ \bar{X}_2 \end{pmatrix}$$

Then the multivariate CLT states that

$$\sqrt{n}(\bar{\mathbf{X}}_n - \boldsymbol{\mu}) \xrightarrow{d} N(\mathbf{0}, \Sigma)$$

We now define the following scalar function of a vector,  $g : \mathbb{R}^2 \mapsto \mathbb{R}$ :

$$g \left[ \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right] := \frac{y_1}{y_2}$$

Which has gradient given by

$$\nabla_{\mathbf{y}} g(\mathbf{y}) = \begin{pmatrix} \frac{\partial g}{\partial y_1} \\ \frac{\partial g}{\partial y_2} \end{pmatrix}$$

Denoting the gradient evaluated at the mean,  $\nabla_{\mathbf{y}} g(\mathbf{y})|_{\mathbf{y}=\boldsymbol{\mu}}$  as  $\nabla_{\boldsymbol{\mu}}$ , the multivariate Delta Method states that

$$\sqrt{n} (g(\bar{\mathbf{X}}_n) - g(\boldsymbol{\mu})) \xrightarrow{d} N(\mathbf{0}, \nabla_{\boldsymbol{\mu}}^T \Sigma \nabla_{\boldsymbol{\mu}})$$

if  $g(\cdot)$  is smooth and differentiable.

The partial derivatives of the scalar function  $g(\cdot)$  with respect to  $y_1$  and  $y_2$  are

$$\frac{\partial g}{\partial x_1} = \frac{1}{x_2}, \quad \frac{\partial g}{\partial x_2} = \frac{-x_1}{x_2^2}$$

Evaluating this at  $\mathbf{y} = \boldsymbol{\mu}$  we have that

$$\nabla_{\boldsymbol{\mu}} = \begin{pmatrix} \frac{1}{\mu_2} \\ \frac{-\mu_1}{\mu_2^2} \end{pmatrix}$$

In order to compute the asymptotic variance of  $\sqrt{n} (g(\bar{\mathbf{X}}_n) - g(\boldsymbol{\mu}))$ , we have to compute

$$\begin{aligned} \nabla_{\boldsymbol{\mu}}^T \Sigma \nabla_{\boldsymbol{\mu}} &= \begin{pmatrix} \frac{1}{\mu_2} & \frac{-\mu_1}{\mu_2^2} \end{pmatrix} \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \begin{pmatrix} \frac{1}{\mu_2} \\ \frac{-\mu_1}{\mu_2^2} \end{pmatrix} \\ &= \begin{pmatrix} \frac{\sigma_{11}}{\mu_2} - \frac{\sigma_{21}\mu_1}{\mu_2^2}, & \frac{\sigma_{12}}{\mu_2} - \frac{\sigma_{22}\mu_1}{\mu_2^2} \end{pmatrix} \begin{pmatrix} \frac{1}{\mu_2} \\ \frac{-\mu_1}{\mu_2^2} \end{pmatrix} \\ &= \frac{1}{\mu_2} \left( \sigma_{11}\mu_2 - \frac{\sigma_{21}\mu_1}{\mu_2^2} \right) - \frac{\mu_1}{\mu_2^2} \left( \frac{\sigma_{12}}{\mu_2} - \frac{\sigma_{22}\mu_1}{\mu_2^2} \right) \\ &= \left( \frac{\sigma_{11}}{\mu_2^2} - \frac{\sigma_{21}\mu_1}{\mu_2^3} \right) - \left( \frac{\mu_1\sigma_{12}}{\mu_2^3} - \frac{\sigma_{22}\mu_1^2}{\mu_2^4} \right) \\ &= \frac{\sigma_{11}\mu_2^2 - \sigma_{21}\mu_1\mu_2 - \mu_1\mu_2\sigma_{12} + \sigma_{22}\mu_1^2}{\mu_2^4} \\ &= \frac{1}{\mu_2^4} (\sigma_{11}\mu_2^2 - (\sigma_{12} + \sigma_{21})\mu_1\mu_2 + \sigma_{22}\mu_1^2) \end{aligned}$$

And we have the limiting distribution

$$\sqrt{n} (\bar{X}_1 / \bar{X}_2 - \mu_1 / \mu_2) \xrightarrow{d} N \left( 0, \frac{1}{\mu_2^4} (\sigma_{11}\mu_2^2 - (\sigma_{12} + \sigma_{21})\mu_1\mu_2 + \sigma_{22}\mu_1^2) \right)$$

Which yields the following result on the limiting distribution of  $Y_n$ :

$$Y_n = \frac{\bar{X}_1}{\bar{X}_2} \xrightarrow{d} N \left( \frac{\mu_1}{\mu_2}, \frac{\sigma^2}{n} \right)$$

where  $\sigma^2 = \frac{1}{\mu_2^4} (\sigma_{11}\mu_2^2 - (\sigma_{12} + \sigma_{21})\mu_1\mu_2 + \sigma_{22}\mu_1^2)$