Class Notes

Estimation in High-Dimensional Spaces - ECON231C

Mauricio Vargas-Estrada
Master in Quantitative Economics
University of California - Los Angeles

1 Markov's Inequality

Being X a random variable such that $X \geq 0$, then:

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}[X]}{t}, \quad \forall t > 0$$

Proof. Whe can rewrite the left-hand side of the inequality using the indicator function:

$$X \geq X\mathbf{1}_{\{X \geq t\}}$$

The left-hand side would be greater when $X < t$ and equal when $X \geq t$. Given that:

$$X \geq X\mathbf{1}_{\{X \geq t\}} \geq t\mathbf{1}_{\{X \geq t\}}$$

In this case, $X\mathbf{1}_{\{X \geq t\}} > t\mathbf{1}_{\{X \geq t\}}$ when $X > t$, and $X\mathbf{1}_{\{X \geq t\}} = t\mathbf{1}_{\{X \geq t\}}$ when $X \leq t$ because $X = t$ or the indicator function is zero.

Taking the expectation of the inequality:

$$\begin{aligned}
\mathbb{E}[X] &\geq \mathbb{E}\left[X\mathbf{1}_{\{X \geq t\}}\right] \geq \mathbb{E}\left[t\mathbf{1}_{\{X \geq t\}}\right] \\
\mathbb{E}[X] &\geq \mathbb{E}\left[X\mathbf{1}_{\{X \geq t\}}\right] \geq t\mathbb{E}\left[\mathbf{1}_{\{X \geq t\}}\right] \\
\frac{\mathbb{E}[X]}{t} &\geq \frac{\mathbb{E}\left[X\mathbf{1}_{\{X \geq t\}}\right]}{t} \geq \mathbb{E}\left[\mathbf{1}_{\{X \geq t\}}\right]
\end{aligned}$$

But $\mathbb{E}\left[\mathbf{1}_{\{X \geq t\}}\right] = \mathbb{P}(X \geq t)$, so:

$$\frac{\mathbb{E}[X]}{t} \geq \mathbb{P}(X \geq t)$$

□

2 Chevyshev's Inequality

Given a random variable X with mean μ and variance σ^2, then:

$$\mathbb{P}\left(|X-\mu|\geq t\right) \leq \frac{\sigma^2}{t^2}, \quad \forall t > 0$$

Proof. We are going to use the fact that a strictly increasing function of a random variable does not change the probability of an event. Let $Y = (|X-\mu|)^2$. Then,

$$\mathbb{P}\left(|X-\mu|^2\geq t^2\right) = \mathbb{P}\left(Y \geq t^2\right)$$

Using Markov's inequality, we have:

$$\mathbb{P}\left(Y \geq t^2\right) \leq \frac{\mathbb{E}\left[Y\right]}{t^2}$$

Given a random variable Z, the variance of Z is $\mathrm{Var}\left(Z\right) = \mathbb{E}\left[(Z-\mathbb{E}\left[Z\right])^2\right]$. Therefore,

$$\begin{aligned}
\mathbb{P}\left((X-\mu)^2 \geq t^2\right) &\leq \frac{\mathbb{E}\left[(X-\mu)^2\right]}{t^2} \\
\mathbb{P}\left((X-\mu)^2 \geq t^2\right) &\leq \frac{\mathrm{Var}\left(X\right)}{t^2} \\
\mathbb{P}\left((X-\mu)^2 \geq t^2\right) &\leq \frac{\sigma^2}{t^2}
\end{aligned}$$

□

3 Weak Law of Large Numbers

Given a collection of i.i.d. random variables $\{X_i\}_{i=1}^n$, with mean μ and variance σ^2.

Defining $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$, the weak law of large numbers states that for any $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left(|\bar{X}_n - \mu|\geq \epsilon\right) = 0$$

or equivalently,

$$\bar{X}_n \xrightarrow{\mathbb{P}} \mu$$

Proof. Calculating the variance of $\bar{X}_n$,

$$\begin{aligned}
\operatorname{Var}\left(\bar{X}_n\right) &= \operatorname{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} \operatorname{Var}\left(X_i\right) \\
&= \frac{1}{n^2} \cdot n\sigma^2 \\
&= \frac{\sigma^2}{n}
\end{aligned}$$

Then, by Chebyshev's inequality,

$$\begin{aligned}
\mathbb{P}\left(|\bar{X}_n - \mu| \geq \epsilon\right) &\leq \frac{\operatorname{Var}\left(\bar{X}_n\right)}{\epsilon^2} \\
&= \frac{\sigma^2}{n\epsilon^2}
\end{aligned}$$

Taking the limit as $n \to \infty$,

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{P}\left(|\bar{X}_n - \mu| \geq \epsilon\right) &\leq \lim_{n\to\infty} \frac{\sigma^2}{n\epsilon^2} \\
\lim_{n\to\infty} \mathbb{P}\left(|\bar{X}_n - \mu| \geq \epsilon\right) &\leq 0
\end{aligned}$$

□

4 Hoeffding's Inequality

If $\{X_i\}_{i=1}^n$ is a random sample from a distribution with mean μ such that, for a number [1] $a > 0$, we have:

$$|X_i - \mu| \leq a, \quad \forall i = 1, 2, \ldots, n$$

Then, for any $t > 0$, the following inequality holds:

$$\mathbb{P}\left(|\bar{X}_n - \mu| \geq t\right) \leq 2\exp\left(-\frac{nt^2}{2a^2}\right)$$

where $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ is the sample mean.

Proof. Lets define $Z_i = X_i - \mu, \forall i = 1, 2, \ldots, n$. Then, we have:

[1] In the field of statistics, it is common to consider bounded random variables, which naturally leads to the assumption that all moments exist. However, in econometrics, it's often more pragmatic to soften this assumption, focusing instead on the existence of only a select subset of moments. This approach allows for greater flexibility in dealing with real-world data, where the behavior of economic variables can't always be neatly bounded, and full moment conditions may not hold.

$$\bar{Z}_n = \frac{1}{n}\sum_{i=1}^{n} Z_i = \bar{X}_n - \mu$$

and,

$$|Z_i| \leq a, \forall i = 1, 2, \ldots, n$$

Consider the events:

$$A = \{\bar{Z}_n \geq t\}$$
$$B = \{\bar{Z}_n \leq -t\}$$

then, the probability of a event $C = \{|\bar{Z}_n| \geq t\}$, can be written as:

$$\mathbb{P}(C) = \mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$$

by the union bound. Now, we can write:

$$\mathbb{P}\left(|\bar{Z}_n| \geq t\right) \leq \mathbb{P}\left(\bar{Z}_n \geq t\right) + \mathbb{P}\left(\bar{Z}_n \leq -t\right)$$
$$\mathbb{P}\left(|\bar{Z}_n| \geq t\right) \leq \mathbb{P}\left(\sum_{i=1}^{n} Z_i \geq nt\right) + \mathbb{P}\left(\sum_{i=1}^{n} Z_i \leq -nt\right)$$

for any $\lambda > 0$, we have [2]:

$$\mathbb{P}\left(|\bar{Z}_n| \geq t\right) \leq \mathbb{P}\left(\lambda\sum_{i=1}^{n} Z_i \geq \lambda nt\right) + \mathbb{P}\left(\lambda\sum_{i=1}^{n} Z_i \leq -\lambda nt\right)$$

and, by Markov's inequality:

$$\mathbb{P}\left(|\bar{Z}_n| \geq t\right) \leq \frac{\mathbb{E}\left[\exp(\lambda\sum_{i=1}^{n} Z_i)\right]}{\exp(\lambda nt)} + \frac{\mathbb{E}\left[\exp(-\lambda\sum_{i=1}^{n} Z_i)\right]}{\exp(-\lambda nt)}$$

since Z_i are independent and identically distributed, we can write:

[2] Given that $f(x) = \lambda x$ is a monotonically increasing function when $\lambda > 0$

$$
\begin{aligned}
\mathbb{P}\left(|\bar{Z}_n| \geq t\right) &\leq \frac{\prod_{i=1}^{n} \mathbb{E}\left[\exp(\lambda Z_i)\right]}{\exp(\lambda n t)} + \frac{\prod_{i=1}^{n} \mathbb{E}\left[\exp(-\lambda Z_i)\right]}{\exp(-\lambda n t)} \\
\mathbb{P}\left(|\bar{Z}_n| \geq t\right) &\leq \frac{\prod_{i=1}^{n} \mathbb{E}\left[\exp(\lambda Z_i)\right]}{\exp(\lambda n t)} + \frac{\prod_{i=1}^{n} \exp(-1)\mathbb{E}\left[\exp(\lambda Z_i)\right]}{\exp(-\lambda n t)} \\
\mathbb{P}\left(|\bar{Z}_n| \geq t\right) &\leq \frac{2\prod_{i=1}^{n} \mathbb{E}\left[\exp(\lambda Z_i)\right]}{\exp(\lambda n t)}
\end{aligned}
$$

Applying the Hoeffding's lemma [3] to the above expression, we have:

$$
\mathbb{P}\left(|\bar{Z}_n| \geq t\right) \leq \frac{2\prod_{i=1}^{n} \exp\left(\frac{\lambda^2 a^2}{2}\right)}{\exp(\lambda n t)}
$$

simplifying the above expression:

$$
\begin{aligned}
\mathbb{P}\left(|\bar{Z}_n| \geq t\right) &\leq \frac{2\exp\left(\frac{n\lambda^2 a^2}{2}\right)}{\exp(\lambda n t)} \\
\mathbb{P}\left(|\bar{Z}_n| \geq t\right) &\leq 2\exp\left(\frac{n\lambda^2 a^2}{2} - \lambda n t\right)
\end{aligned}
$$

Because the above inequality holds for any $\lambda > 0$, we can optimize the right-hand side with respect to λ.

$$
\lambda^* = \arg\min_{\lambda > 0} \left\{ \frac{n\lambda^2 a^2}{2} - \lambda n t \right\}
$$

Calculating the F.O.C. with respect to λ, we get:

$$
n a^2 \lambda^* - n t = 0 \Rightarrow \lambda^* = \frac{t}{a^2}
$$

Substituting λ^* back into the inequality:

$$
\begin{aligned}
\mathbb{P}\left(|\bar{Z}_n| \geq t\right) &\leq 2\exp\left(\frac{n\left(\frac{t}{a^2}\right)^2 a^2}{2} - \frac{t}{a^2} n t\right) \\
\mathbb{P}\left(|\bar{Z}_n| \geq t\right) &\leq 2\exp\left(\frac{n t^2}{2a^2} - \frac{n t^2}{a^2}\right) \\
\mathbb{P}\left(|\bar{Z}_n| \geq t\right) &\leq 2\exp\left(-\frac{n t^2}{2a^2}\right)
\end{aligned}
$$

[3] If X is a random variable such that $X \leq a$, then for any $\lambda > 0$, we have:

$$
\mathbb{E}\left[\exp(\lambda X)\right] \leq \exp\left(\frac{\lambda^2 a^2}{2}\right)
$$

replacing $\bar{Z}_n$ by $\bar{X}_n - \mu$:

$$\mathbb{P}\left(|\bar{X}_n - \mu| \geq t\right) \leq 2\exp\left(-\frac{nt^2}{2a^2}\right)$$

$\square$