Dr. Z.'s Probability Lecture 22 Handout:
The Markov and Chebyshev Inequalities and the Weak Law of Large Numbers

By Doron Zeilberger

Version of Dec. 16, 2017 (Thanks to Ryan Lee.)

Important Inequality (Markov's inequality): If X is a random variable that takes only non-negative values, then for any $a > 0$:

$$P\{X \geq a\} \leq \frac{E[X]}{a} \quad .$$

Note: Because Markov's inequality is so general, it is, usually, very uninformative. It gives you a guaranteed upper bound for $P[X \geq a]$ but usually the actual value is very far below it.

First Proof of Markov's Inequality:

Discrete Case:

If X is discrete random variable, let $p(x)$ be its probability mass function. By definition (since $p(x)$ is 0 for negative x)

$$E[X] = \sum_{0 \leq x < \infty} x\,p(x) \quad .$$

Let's break-up the sum:

$$E[X] = \sum_{0 \leq x < a} x\,p(x) \,+ \sum_{a \leq x < \infty} x\,p(x)$$

$$\geq \sum_{a \leq x \leq \infty} xp(x) \,\geq\, a \sum_{a \leq x \leq \infty} p(x) \,=\, a\,P\{X \geq a\} \quad .$$

Hence $E[X] \geq a\,P\{X \geq a\}$ that implies $P\{X \geq a\} \leq \frac{E[X]}{a}$.

Continuous Case: If X is a continuous random variable, let $f(x)$ be its probability density function. By definition (since $f(x)$ is 0 for negative x), we have

$$E[X] = \int_0^\infty x f(x)\,dx \quad .$$

Let's break-up the integral

$$E[X] = \int_0^a x f(x)\,dx \,+ \int_a^\infty x f(x)\,dx$$

$$\geq \int_a^\infty x f(x)\,dx \,\geq\, a \int_a^\infty f(x)\,dx = a\,P\{X \geq a\} \quad .$$

Hence $E[X] \,\geq\, a\,P\{X \geq a\}$ that implies $P\{X \geq a\} \,\leq\, \frac{E[X]}{a}$.

The Book Proof: The book proof is shorter, and contains both cases (including more general types, where X is a mixture of continuous and discrete) but is a bit abstract.

For $a > 0$, let

$$I = \begin{cases} 1 \quad , & if \quad X \geq a\,; \\ 0 \quad , & otherwise. \end{cases} \quad .$$

Obviously (since $X \geq 0$)

$$I \leq \frac{X}{a} \quad .$$

(When $X < a$ the left side is 0 and the right side is non-negative, when $X \geq a$ the left side is 1 and the right side is ≥ 1 (and usually, > 1).

Taking expectation, we have

$$E[I] \leq E[\frac{X}{a}] = \frac{E[X]}{a} \quad .$$

But $E[I] = 0 \cdot P[X < a] + 1 \cdot P[X \geq a] = P[X \geq a]$, so the inequality follows.

Problem 22.1: Let X be the random variable whose probability density function is

$$f(x) = \begin{cases} \frac{1}{(\ln 2)\, x}\,, & if \quad 1 \leq x \leq 2\,; \\ 0\,, & otherwise. \end{cases} \quad .$$

Verify Markov's inequality for $a = 1.5$, by finding $P\{X \geq 1.5\}$ and $E[X]/1.5$.

Sol. of 22.1: We first need the expectation

$$E[X] = \int_1^2 x \frac{1}{(\ln 2)\, x}\, dx = \int_1^2 \frac{1}{\ln 2}\, dx = \frac{1}{\ln 2} \quad .$$

Also

$$P\{X \geq 1.5\} = \int_{1.5}^2 \frac{1}{\ln 2} \frac{1}{x}\, dx = \frac{1}{\ln 2} \left(\ln x \Big|_{1.5}^2 \right) = \frac{1}{\ln 2}(\ln 2 - \ln 1.5) = \frac{\ln 2 - \ln 1.5}{\ln 2} = 0.4150374993\ldots \quad .$$

On the other hand

$$\frac{E[X]}{1.5} = \frac{1}{1.5 \cdot \ln 2} = 0.9617966940\ldots \quad .$$

Since $0.4150374993 < 0.9617966940\ldots$, we verified Markov's inequality for this case.

An even more important inequality: Chebyshev's inequality:

If X is random variable with mean μ and variance σ^2, then for any value $k > 0$,

$$P\{|X - \mu| \geq k\} \leq \frac{\sigma^2}{k^2} \quad .$$

Note: The inequality implies that the probability of **not** being $\pm k$ from the average is less than $\frac{\sigma^2}{k^2}$. Hence the further you go away from the average, the less likely it is to be there. Once again, in many cases the inequality, while obviously correct, gives a way too conservative bound. For example if $k < \sigma$ it is not at all informative, since probabilities are always ≤ 1. But for $k > \sigma$ it starts telling you something. For example the probabilities that you are

- further than two standard-deviations from the mean is $\leq \frac{1}{4}$,

(just for comparison, for a Normal distribution, the actual value is is $2\,\Phi(-2) \,=\, 0.04550026388\ldots$),

- further than three standard-deviations from the mean is $\leq \frac{1}{9}$,

(just for comparison, for a Normal distribution, the actual value is is $2\,\Phi(-3) \,=\, 0.002699796402\ldots$),

- further than five standard-deviations from the mean is $\leq \frac{1}{25}$,

(just for comparison, for a Normal distribution, the actual value is is $2\,\Phi(-5) \,=\, 0.0000005736789990\ldots$).

Proof of Chebyshev's Inequality: $(X-\mu)^2$ is a random variable that is always non-negative, hence Markov's inequality (with $a = k^2$) applies

$$P\{(X-\mu)^2 \,\geq\, k^2\} \,\leq\, \frac{E[(X-\mu)^2]}{k^2} \quad .$$

But, by definition $E[(X-\mu)^2] = \sigma^2$, and saying that $(X-\mu)^2 \,\geq\, k^2$ is the same thing as saying that $|X-\mu| \,\geq\, k$, we are done.

Problem 22.2: Find the mean, μ and standard deviation, σ of the random variable X whose density function, $f(x)$ is given by

$$f(x) \,=\, \begin{cases} \frac{x}{4}\,, if \quad 1 \leq x \leq 3\,; \\ 0\,, otherwise. \end{cases} \quad ,$$

and verify Chebyshev's inequality for $k = 3\sigma$.

Sol. of 22.2:

$$\mu \,=\, \int_1^3 \frac{x^2}{4} \,=\, \frac{13}{6} = 2.166666666\ldots \quad ,$$

$$E[X^2] \,=\, \int_1^3 \frac{x^3}{4} \,=\, 5 \quad .$$

Hence

$$\sigma^2 \,=\, E[X^2] - E[X]^2 \,=\, 5 - (\frac{13}{6})^2 \,=\, \frac{11}{36} \quad .$$

The right side of Chebyshev's inequality is $\sigma^2/(3\sigma)^2 = \frac{1}{9}$. Now $k = 3\sqrt{\frac{11}{36}} = \sqrt{11}/2 = 1.658312395$. We have

$$P\{|X-\mu| \geq k\} = \int_1^{\mu-k} f(x)\,dx \,+\, \int_{\mu+k}^3 f(x)\,dx \quad ,$$

Since $\mu - k \,<\, 1$ and $\mu + k \,>\, 3$ **both** integrals are zero, Since $0 \,<\, \frac{1}{9}$, Chebyshev's inequality is confirmed in this case.

Problem 22.3: Suppose that it is known that the amount of uranium dug in an uranium mine during one week is a random variable with mean 20 kg.

(i) What can be said about the probability that the week's production will exceed 24 kg?

(ii) If the standard deviation of the amount of uranium dug in the mine equals 4 kg, what can be said about the week's production will be between 8 and 32 kg?

Sol. to 22.3: For (i) we are supposed to use Markov, for (ii) Chebyshev.

For **(i)**, $E[X] = \mu = 20$ and $a = 24$. By Markov

$$P\{X \geq 24\} \leq \frac{20}{24} = \frac{5}{6} \quad .$$

Ans. to 22.3(i): The probability that the week's production will exceed 24 kilograms is at at most $\frac{5}{6}$.

For **(ii)**, we need Chebyshev. First let's rearrange:

$$P\{8 \leq X \leq 32\} = P\{8-20 \leq X-20 \leq 32-20\} = P\{-12 \leq X-20 \leq 12\} = P\{|X-20| \leq 12\} \quad .$$

We have: $\mu = 20$ and $\sigma = 4$. By Chebyshev (with $k = 12$)

$$P\{|X-20| \geq 12\} \leq \frac{\sigma^2}{k^2} = \frac{4^2}{12^2} = \frac{1}{9} \quad .$$

Hence

$$P\{|X-20| \leq 12\} \geq 1 - \frac{1}{9} = \frac{8}{9} \quad .$$

Ans. to 22.3(ii): The probability that the week's production will be between 8 and 32 kg is at least $\frac{8}{9}$.

A famous and important theorem (The Weak Law of Large Numbers):

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed random variables, each having a finite mean $E[X_i] = \mu$. Then for any $\epsilon > 0$,

$$\lim_{n \to \infty} P\{ \left| \frac{X_1 + \ldots + X_n}{n} - \mu \right| \geq \epsilon \} = 0 \quad .$$

Proof under the assumption of finite variance. We are assuming that $Var(X_i) = \sigma^2$ is finite.

By the linearity of expectation

$$E\left[\frac{X_1 + \ldots + X_n}{n}\right] = \frac{1}{n}\left(E[X_1] + \ldots + E[X_n]\right) = \frac{1}{n}(n\mu) = \mu \quad .$$

By the linearity of variance for **independent** random variables

$$Var(\,X_1 + \ldots + X_n\,) = Var(X_1) + \ldots + Var(X_n) \;=\; n\sigma^2 \quad .$$

Since $Var(aY) \;=\; a^2\, Var(Y)$ we have

$$Var(\,\frac{X_1 + \ldots + X_n}{n}\,) = (\frac{1}{n})^2\, n\, \sigma^2 = \frac{\sigma^2}{n} \quad .$$

Thanks to Chebyshev

$$P\{\,|\frac{X_1 + \ldots + X_n}{n} \,-\, \mu| \,\geq\, \epsilon\} \leq \frac{\sigma^2}{n\epsilon^2} \quad .$$

Since ϵ, and hence $\frac{\sigma^2}{\epsilon^2}$ is a **constant**, the theorem follows by letting n go to infinity.