Worksheet 19 (Solutions)

1. Let $Z \sim N(0,1)$. It can be shown that $\mathbb{E}|Z| = \sqrt{2/\pi}$. Use Markov's inequality to bound the probabilities: (a) $\mathbb{P}[|Z| > 1.28]$, (b) $\mathbb{P}[|Z| > 1.96]$, (c) $\mathbb{P}[|Z| > 2.58]$, and (d) $\mathbb{P}[|Z| > 3.89]$. Compare these to the exact quantities on the handout.

Solution:

This is quite straightforward:

$$\mathbb{P}[|Z| \geq 1.28] \leq \frac{\sqrt{2/\pi}}{1.28} \approx 0.623$$
$$\mathbb{P}[|Z| \geq 1.96] \leq \frac{\sqrt{2/\pi}}{1.96} \approx 0.407$$
$$\mathbb{P}[|Z| \geq 2.58] \leq \frac{\sqrt{2/\pi}}{2.58} \approx 0.309$$
$$\mathbb{P}[|Z| \geq 3.89] \leq \frac{\sqrt{2/\pi}}{3.89} \approx 0.205$$

So, these are decreasing much (much) slower than the exact values.

2. Chebychev's inequality (see the reference sheet) can be derived directly from Markov's inequality. Let X be a random variable with mean μ and variance σ^2. Define $Y = (X - \mathbb{E}X)^2$ and apply Markov's inequality with $X \to Y$ and $a \to a^2$ (remember, a can be any positive constant so we can replace it with a squared version of itself if we do so on both sides). Plug the value of Y back in, use the definition of variance, and simplify to derive Chebychev's inequality

Solution: Apply Markov's Inequality with $X \to Y$ and $a \to a^2$ and then plugging back the value for Y yields:

$$\mathbb{P}[|Y| > a^2] \leq \frac{\mathbb{E}|Y|}{a^2}$$
$$\mathbb{P}[(X - \mathbb{E}X)^2 > a^2] \leq \frac{\mathbb{E}(X - \mathbb{E}X)^2}{a^2}$$
$$\mathbb{P}[(X - \mu)^2 > a^2] \leq \frac{\sigma^2}{a^2}$$

Taking the square root of both sides inside the probability gives:

$$\mathbb{P}[|X - \mu| > a] \leq \frac{\sigma^2}{a^2}$$

And that's all we need for Chebychev's inequality.

3. Let $Z \sim N(0,1)$. Use Chebychev's inequality to bound the probabilities: (a) $\mathbb{P}[|Z| > 1.28]$, (b) $\mathbb{P}[|Z| > 1.96]$, (c) $\mathbb{P}[|Z| > 2.58]$, (d)

$\mathbb{P}[|Z| > 3.89]$. Compare these to the previous results. Which ones are tighter?

Solution:

This is quite straightforward as well:

$$\mathbb{P}[|Z| \geq 1.28] \leq \frac{1}{1.28^2} \approx 0.610$$
$$\mathbb{P}[|Z| \geq 1.96] \leq \frac{1}{1.96^2} \approx 0.260$$
$$\mathbb{P}[|Z| \geq 2.58] \leq \frac{1}{2.58^2} \approx 0.150$$
$$\mathbb{P}[|Z| \geq 3.89] \leq \frac{1}{3.89^2} \approx 0.066$$

These are tighter bounds than Markov gives, particularly for the last two. But, they are still quite a ways away from the exact values.

4. Chernoff's inequality (see the reference sheet) can also be derived directly from Markov's inequality. Let X be a random variable with a well-defined moment generating function. Apply Markov's inequality with $|X| \to e^{tX}$ (the new value is also positive, so no need for absolute value) and $a \to e^{ta}$. Simplify the part inside of the probability on the left-hand side to derive Chernoff's inequality.

Solution: Apply Markov's Inequality with $|X| \to e^{tX}$ and $a \to e^{ta}$ yields:

$$\mathbb{P}[e^{tX} > e^{ta}] \leq \frac{\mathbb{E}e^{tX}}{e^{ta}}$$

Taking the log of both sides in the interior of the probabilty gives:

$$\mathbb{P}[tX > ta] \leq \frac{\mathbb{E}e^{tX}}{e^{ta}}$$
$$\mathbb{P}[X > a] \leq \frac{\mathbb{E}e^{tX}}{e^{ta}}$$

And that's Chernoff's inequality as written on the worksheet.

5. Chernoff's inequality has an extra term in it, the t, that provides a whole family of bounds for a given value of a. The tightest bound depends on the distribution. Let $Z \sim N(0,1)$. Using the moment generating function, what value of t provides the tightest bound on $\mathbb{E}[Z \geq a]$?

Solution: Plugging in the moment generating function, we have the following bound of Z:

$$\begin{aligned}\mathbb{P}[X > a] &\leq \frac{\mathbb{E}e^{tX}}{e^{ta}} \\ &\leq \frac{e^{\frac{1}{2}t^2}}{e^{ta}} = e^{\frac{1}{2}t^2 - ta}\end{aligned}$$

The quantity on the right will be minimized by the value in the exponent $(\frac{1}{2}t^2 - ta)$ is minimized. This is a quadratic polynomial; taking the derivative and setting it equal to 0 gives $t = a$ at the minimum, which has the following final bound on the tail probability from Chernoff's inequality for a standard normal Z:

$$\mathbb{P}[X > a] \leq e^{-\frac{1}{2}a^2}.$$

That turns out to be the correct limiting distribution of the tail of a normal.

6. Let $Z \sim N(0, 1)$. Use Chernoff's inequality (and the tightest value of t from the previous question) to compute bounds on the following: (a) $\mathbb{P}[|Z| > 1.28]$, (b) $\mathbb{P}[|Z| > 1.96]$, (c) $\mathbb{P}[|Z| > 2.58]$, and (d) $\mathbb{P}[|Z| > 3.89]$. Note that due to the symmetry of the normal distribution, you can double the probability that Z is larger than some a to get the probability that $|Z|$ is larger than a. You should notice an interesting pattern relative to the other bounds that we have.

Solution:

This is quite straightforward as well:

$$\begin{aligned}
\mathbb{P}[|Z| \geq 1.28] &\leq 2 \times e^{-\frac{1}{2}(1.28)^2} \approx 0.882 \\
\mathbb{P}[|Z| \geq 1.96] &\leq 2 \times e^{-\frac{1}{2}(1.96)^2} \approx 0.293 \\
\mathbb{P}[|Z| \geq 2.58] &\leq 2 \times e^{-\frac{1}{2}(2.58)^2} \approx 0.072 \\
\mathbb{P}[|Z| \geq 2.58] &\leq 2 \times e^{-\frac{1}{2}(3.89)^2} \approx 0.001
\end{aligned}$$

The first is much more than even the basic Markov bound. The second is better than the Markov bound and just a little worse than Chebyshev. The third, and particularly the fourth, are much better than the previous bounds.

7. (Weak Law of Large Numbers) Let's finish with a result that shows the power of these tail inequalities for establishing theoretical results. Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables that come from a distribution with finite mean μ and finite variance σ^2. For any positive n, define the sample mean to be:

$$\bar{X}_n = \frac{X_1 + \cdots + X_n}{n}$$

Then, for any $\epsilon > 0$:

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

Prove that this is true using Chebyshev's inequality. Hint: Compute the mean and variance of $\bar{X}_n$ and then just apply the theorem as-is.

Solution: The mean and variance of $\bar{X}_n$ are:

$$\mathbb{E}\bar{X}_n = \frac{1}{n}\sum_i \mathbb{E}X_i = \frac{1}{n} \times n \cdot \mu = \mu$$
$$Var\bar{X}_n = \frac{1}{n^2}\sum_i Var(X_i) = \frac{1}{n^2} \times n \cdot \sigma^2 = \sigma^2/n$$

Plugging these into Chebyshev's inequality with $a = \epsilon$, we get:

$$\mathbb{P}[|\bar{X}_n - \mu| > \epsilon] \leq \frac{\sigma^2/n}{\epsilon^2}$$

Notice that the limit of the right-hand side will go towards zero for a sufficently large n, and therefore we have proved the Weak Law of Large Numbers as stated in the question.