Mathematics 1550H – Probability I: Introduction to Probability
TRENT UNIVERSITY, Summer 2023 (S62)

Solutions to Assignment #5
Chebyshev's Inequality

Suppose the random variable Y counts the number of heads in 10 tosses of a biased coin with $P(H) = 0.6$ and $P(T) = 0.4$ on each toss. Y then has a binomial distribution: it can take on the integer values 0 through 10, each with with probability $P(Y = y) = \binom{10}{y} \cdot 0.6^y \cdot 0.4^{10-y}$. It also has expected value $E(X) = 10 \cdot 0.6 = 6$ and variance $V(X) = 10 \cdot 0.6 \cdot 0.4 = 2.4$.

1. Work out $P(|Y - 6| > 3)$ as precisely as you can. *[3]*

SOLUTION. Recall that from our handy cheat sheet of common distributions [1] that a binomial distribution with $n = 10$ and $p = 0.6$ (and hence $q = 0.4$) has as its probability distribution function $m(k) = P(Y = k) = \binom{10}{k} \cdot 0.6^k \cdot 0.4^{10-k}$, where k could be any of 0, 1, ... , 10. It follows that:

$$\begin{aligned} P(|Y-6|>3) &= P(Y-6<-3)+P(Y-6>3) = P(Y<3)+P(Y>9) \\ &= P(Y=0)+P(Y=1)+P(Y=2)+P(Y=10) \\ &= \binom{10}{0}\cdot 0.6^0\cdot 0.4^{10-0} + \binom{10}{1}\cdot 0.6^1\cdot 0.4^{10-1} \\ &\qquad + \binom{10}{2}\cdot 0.6^2\cdot 0.4^{10-2} + \binom{10}{10}\cdot 0.6^10\cdot 0.4^{10-10} \\ &= 1\cdot 1\cdot 0.0001048576 + 10\cdot 0.6\cdot 0.000262144 \\ &\qquad + 45\cdot 0.36\cdot 0.00065536 + 1\cdot 0.0060466176\cdot 1 \\ &\approx 0.0001048576 + 0.001572864 + 0.010616832 + 0.0060466176 \\ &= 0.0183411712 \end{aligned}$$

Ten decimal places is likely to be good enough, and likely overkill, for most purposes. $\square$

Chebyshev's Inequality, which we'll see in class later on, lets us estimate probabilities involving random variables knowing nothing about the distribution involved except for its expected value and standard deviation. Mind you, these estimates are usually pretty crude, but they are pretty easy to get.

There are several ways to state this inequality; one common version states that a random variable X, which could be discrete or continuous, with expected value $\mu = E(X)$ and standard deviation $\sigma = \sqrt{V(X)}$ must satisfy

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

for any real number $k > 0$. Technically, this form of Chebyshev's Inequality gives us upper bounds for certain probabilities.

If the mean and/or the standard deviation are not defined or are unknown for X, Chebyshev's Inequality isn't useful. Note also that it is of interest only if $k > 1$, since if $k \leq 1$, then $\frac{1}{k^2} \geq 1$ and any probability must be ≤ 1.

2. Use Chebyshev's Inequality to find an upper bound for the probability $P(|Y-6|>3)$ in question **1**. How close is this upper bound to the precise value? *[3]*

SOLUTION. Our first little problem is that $P(|Y-6|>3)$ is not quite like the probability in Chebyshev's Inequality, $P(|X-\mu| \geq k\sigma)$, because $>$ is not quite the same as $\geq$. In this case, it is easy enough to work around the problem: because Y only takes on integer values, $P(|Y-6|>3) = P(|Y-6| \geq 4)$. We will apply Chebyshev's Inequality to the latter form.

Our second little problem is that while $6 = E(Y)$, we need to break up the 4 into $k\sigma$, where $\sigma^2 = V(Y) = 2.4$. It follows from the last that $\sigma = \sqrt{V(X)} = \sqrt{2.4}$. If $k\sigma = 4$, then $k = 4/\sigma = 4/\sqrt{2.4}$. Thus, using Chebyshev's Inequality:

$$\begin{aligned} P(|Y-6|>3) = P(|Y-6| \geq 4) &= P\left(|Y-6| \geq \frac{4}{\sqrt{2.4}} \cdot \sqrt{2.4}\right) \\ &\leq \frac{1}{\left(4/\sqrt{2.4}\right)^2} = \frac{2.4}{16} = 0.15 \end{aligned}$$

This is larger than the pretty precise value of 0.0183411712 for $P(|Y-6|>3)$s by a factor of a bit more than 8. □

3. Either find

i. a probability density function $f(x)$ for a continuous random variable such that a continuous random variable X with this density will have $\mu = 0$, $\sigma = 1$, and satisfy $P(|X| \geq 2) = \frac{1}{4}$, or

ii. a probability distribution function $m(x)$ for a discrete random variable X such that a discrete random variable X with this distribution function will have $\mu = 0$, $\sigma = 1$, and satisfy $P(|X| \geq 2) = \frac{1}{4}$. *[4]*

SOLUTION. Trying to find an example all by yourself is likely to be, well, trying. Fortunately, you are allowed to look things up on the assignments. Chebyshev's Inequality has its own article [2] on Wikipedia, which actually gives an example of alternative *ii*. A simplified and cleaned up version of this example follows:

Suppose the discrete random variable X returns the values -2, 0, and 2, with the following probabilities: $P(X=-2) = P(X=2) = \frac{1}{8} = 0.125$ and $P(X=0) = \frac{3}{4} = 0.75$. Then the expected value of X is $\mu = E(X) = 0$ and the variance of X is $\sigma^2 = V(X) = 1$, so it has standard deviation $\sigma = 1$. (We leave it to you check these numbers.) Then

$$P(|X-\mu| \geq 2\sigma) = P(|X| \geq 2) = P(X=-2) + P(X=2) = \frac{1}{8} + \frac{1}{8} = \frac{1}{4} = \frac{1}{2^2},$$

as desired. □

NOTE: The point of **3** is that Chebyshev's Inequality can't really be improved upon, short of having additional information about the distribution of the random variable X.

REFERENCES

1. *Some Common Probability Distributions – The Short Form*, MATH 1550H handout in the lecture of 2023-07-12.

2. *Chebyshev's inequality*, *Wikipedia* article, accessed on 2023-07-21 via the URL: `https://en.wikipedia.org/wiki/Chebyshev's_inequality`, .