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

Solutions to Quiz #10
Counting Heads Again

Instructions: Do all of the following problems. Please show all your work.

The discrete random variable X counts the number of heads that come up in 50 tosses of a fair coin.

1. Compute $P(23 \le X \le 27)$ as precisely as you can. *[1]*

NOTE: The solutions to Quiz #8 have an example of computing a probability for a binomial distribution using SageMath that you could modify to compute the one asked for here.

SOLUTION. Since tossing a fair coin and counting the number of heads that come up on one toss is a Bernoulli trial with a probability $p = \frac{1}{2} = 0.5$ of success and $q = 1-p = 1-\frac{1}{2} = \frac{1}{2} = 0.5$ of failure, our random variable X, which counts the number of heads in 50 tosses of the coin, has a binomial distribution with $n = 50$ and $p = q = \frac{1}{2} = 0.5$. Thus the probability distribution function of X is $m(k) = P(X = k) = \binom{50}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{50-k} = \binom{50}{k}\left(\frac{1}{2}\right)^{50}$, where k could be any integer with $0 \le k \le 50$. It follows that:

$$P(23 \le X \le 27) = \sum_{k=23}^{27} P(X = k) = \sum_{k=23}^{27} \binom{50}{k}\left(\frac{1}{2}\right)^{50} = \frac{1}{2^{50}} \sum_{k=23}^{27} \binom{50}{k}$$

As implicitly suggested in the note, to actually compute this we use SageMath, adapting the commmands given in the solutions to Assignment #8:

```
In [1]: var("k")
        sum(binomial(50,k),k,23,27)/2^50
Out[1]: 36599652077597/70368744177664

In [2]: N(36599652077597/70368744177664)
Out[2]: 0.520112338301672
```

Thus $P(23 \le X \le 27) = \dfrac{36599652077597}{70368744177664} \approx 0.5201$. $\square$

2. Use Chebyshev's Inequality to find a lower bound for $P(23 \le X \le 27)$. How close is this lower bound to the actual value? *[1]*

SOLUTION. Observe that $P(23 \le X \le 27) = P(|X - 25| \le 2) = 1 - P(|X - 25| > 2) = 1 - P(|X - 25| \ge 3)$, the last equality following because X can only have integer values. $P(|X - 25| \ge 3)$ is something we can apply Chebyshev's Inequality to since $\mu = E(X) = 25$.

A binomial dsitribution with $n = 50$ and $p = q = \frac{1}{2} = 0.5$ has an expected value of $\mu = E(X) = 50 \cdot \frac{1}{2} = 25$ and a variance of $\sigma^2 = V(X) = 50 \cdot \frac{1}{2} \cdot \frac{1}{2} = 12.5$, and hence a standard deviation of $\sigma = \sqrt{12.5} \approx 3.5355$. It follows that if we write $3 = k\sigma$ in order to apply Chebyshev's Inequality, then $k = \frac{3}{\sigma} \approx \frac{3}{3.5355} \approx 0.8485$. Applying the inequality, we get:

$$P\left(|X - 25| \geq 3\right) \approx P\left(|X - 25| \geq 0.8485 \cdot 3.5355\right) \leq \frac{1}{0.8485^2} \approx 1.3890$$

It follows, in turn, that:

$$P(23 \leq X \leq 27) = 1 - P\left(|X - 25| \geq 3\right) \geq 1 - 1.3890 = -0.3890$$

Since every probability is at least 0, this tells us nothing; that is, Chebushev's Inequality is not actually useful here. □

3. Use the standard normal distribution to approximate $P(23 \leq X \leq 27)$. How close is this approximation to the actual value? *[2]*

Solution. Recall that X counts the number of heads that come up in 50 tosses of a fair coin. Let X_k count the number of heads that come on the kth toss of the fifty. Each X_k is then a Bernoulli trial with probability of success $p = \frac{1}{2}$ and probability of failure $q = \frac{1}{2}$, and hence has expected value $\mu = E\left(X_k\right) = \frac{1}{2}$ and variance $\sigma^2 = V\left(X_k\right) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$, giving it a standard deviation of $\sigma = \sqrt{\frac{1}{4}} = \frac{1}{2}$. Note that $X = X_1 + X_2 + \cdots + X_{50}$ and let

$$S_{50}^* = \frac{X_1 + \cdots + X_{50} - 50\mu}{\sigma\sqrt{50}} = \frac{X - 50 \cdot \frac{1}{2}}{\frac{1}{2} \cdot \sqrt{50}} \approx \frac{X - 25}{3.5355}\,.$$

Then S_{50}^* has an approximately standard normal distribution. It follows that:

$$\begin{aligned} P(23 \leq X \leq 27) &= P(-2 \leq X - 25 \leq 2) = P\left(\frac{-2}{3.5355} \leq \frac{X - 25}{3.5355} \leq \frac{2}{3.5355}\right) \\ &\approx P\left(-0.5657 \leq S_{50}^* \leq 0.5657\right) \approx P(-0.56 \leq Z \leq 0.56) \\ &= P(Z \leq 0.56) - P(\leq -0.56) \approx 0.7123 - 0.2877 = 0.4246 \end{aligned}$$

This value is not particularly close to the more precise 0.5201 obtained in **1** – it's barely within 0.1 – but it is way better than the useless result we got from Chebyshev's Inequality in **2**. □

4. Toss an actual coin 50 times and record the sequence of heads and tails.

a. Give the sequence of heads and tails that you recorded. How many heads came up in the 50 tosses you made? *[0.5]*

b. Do you think the outcome of your 50 tosses support the hypothesis that the coin is more-or-less fair? *[0.5]*

NOTE: Yes, the marks in question **4** ~~are~~ ought to be gift marks. :-)

SOLUTION. **a.** Here is the outcome of your instructor tossing a loonie 50 times:

`TTTTTTTHTHHHHHTTHHHTTHHHHHHHTHHTTHHTHTHTTTHHHHTHTT`

This sequence of 50 tosses has 27 heads and 23 tails. □

b. In this case, yes. The number of heads is within the range that you would expect it to be about half the time according to the solution to question **1**. □

NOTE: The presence of longish runs of heads and of tails in the 50 tosses also tends to support the hypothesis that the randomness is genuine, and not just some bored human typing in a sequence of heads and tails and trying to make it look random.