Math 10B Probability Worksheet 4

1. There are n people who each have their own hat. You take all the hats and randomly rearrange them. Let the random variable X be the number of people who get their own hat back. What is $E[X]$?

   Let X_i be the random variable that is 1 if person i gets back their own hat and 0 otherwise. Since each person is equally likely to get person i's hat, $E[X_i] = P(X_i = 1) = \frac{1}{n}$. Since $X = \sum_{i=1}^{n} X_i$, by linearity of expectation we have

   $$E[X] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \frac{1}{n} = 1.$$

2. Consider the scenario described in problem (1) when there are just two people. What is $\mathrm{Var}[X]$?

   If there are only two people then either both get back their own hat or neither gets back their own hat. And each of these two possibilities happens with probability 1/2. So we can calculate $E[X^2]$ as follows:

   $$E[X^2] = 0 \cdot \frac{1}{2} + 2^2 \cdot \frac{1}{2} = 2.$$

   Therefore

   $$\mathrm{Var}[X] = E[X^2] - E[X]^2 = 2 - 1^2 = 1.$$

3. If X is a random variable and $\mathrm{Var}[X] = 0$, what can you say about X?

   X must be constant.

4. Suppose X is a nonnegative random variable and a is a positive number. Show that $P(X \geq a) \leq \frac{E[X]}{a}$.

Let R denote the range of X. Then we have

$$\begin{aligned}
E[X] &= \sum_{k \in R} kP(X = k) \\
&= \sum_{k \in R:\, k < a} kP(X = k) + \sum_{k \in R:\, k \geq a} kP(X = k) \\
&\geq 0 + \sum_{k \in R:\, k \geq a} kP(X = k) && \text{since } X \text{ is nonnegative} \\
&\geq \sum_{k \in R:\, k \geq a} aP(X = k) \\
&= a \sum_{k \in R:\, k \geq a} P(X = k) \\
&= aP(X \geq a).
\end{aligned}$$

Therefore $E[X] \geq aP(X \geq a)$ so dividing both sides by a gives the desired inequality.
By the way, this is called Markov's inequality and is a surprisingly useful tool in probability.

5. Challenge Question: Show that if X is a random variable with $E[X] = \mu$ and $\text{Var}[X] = \sigma^2$ then for any $k > 0$, $P(|X - \mu| > k\sigma) \leq \frac{1}{k^2}$. [Hint: use the result of the previous problem applied to the random variable $(X - \mu)^2$.]

Consider the random variable $Y = (X - \mu)^2$. Note that Y is nonnegative and $E[Y] = \text{Var}[X] = \sigma^2$. Applying the result from the previous problem to Y and $k^2\sigma^2$ gives

$$P(Y \geq k^2\sigma^2) \leq \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2}.$$

But $Y \geq k^2\sigma^2$ if and only if $|X - \mu| \geq k\sigma$ so we are done.
This inequality is called Chebyshev's inequality and like Markov's inequality is very useful in probability, with numerous applications in math, computer science and other fields.

6. Suppose you roll 20 fair 6-sided dice. Let the random variable X be the sum of the rolls.

   (a) What is $E[X]$?

   For each $i \leq 20$ let the random variable X_i be the value of the i^{th} die. Then $X = X_1 + X_2 + \ldots + X_{20}$. So by linearity of expectation:

   $$E[X] = E[X_1] + E[X_2] + \ldots + E[X_{20}].$$

Now we directly calculate $E[X_i]$:

$$E[X_i] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{7}{2}.$$

Therefore $E[X] = 20 \cdot \frac{7}{2} = 70$.

(b) What is $\text{Var}[X]$?

Note that the random variables X_i from the solution to part (a) are actually independent random variables. Therefore

$$\text{Var}[X] = \text{Var}[X_1] + \text{Var}[X_2] + \ldots + \text{Var}[X_{20}].$$

Now we calculate $\text{Var}[X_i]$:

$$\begin{aligned}
\text{Var}[X_i] &= E[X_i^2] - E[X_i]^2 \\
&= \left(1 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 9 \cdot \frac{1}{6} + 16 \cdot \frac{1}{6} + 25 \cdot \frac{1}{6} + 36 \cdot \frac{1}{6}\right) - \left(\frac{7}{2}\right)^2 \\
&= \frac{91}{6} - \frac{49}{4} \\
&= \frac{35}{12}.
\end{aligned}$$

Therefore $\text{Var}[X] = 20 \cdot \frac{35}{12} = \frac{175}{3}$.