MAT344 Lecture 20

2019/July/25

1 Announcements

2 This week

This week, we are talking about

1. Probability

3 Recap

Last time we talked about

1. Probability
2. Conditional Probability

4 Bernoulli Trials

Example 4.1 (Example 10.11. in [KT17]). *When a die is rolled, let's say that we have a success if the result is a two or a five. The probability p of success is then $\frac{2}{6} = \frac{1}{3}$ and the probability of failure is $\frac{2}{3}$. If the die is rolled ten times in succession, then the probability that we get exactly four successes can be computed by the formula*

$$\binom{10}{4} \left(\frac{1}{3}\right)^4 \left(\frac{2}{3}\right)^6,$$

since four of the ten rolls can be chosen in $\binom{10}{4}$ ways, and the probability of the four successes and six failures in the chosen order is always $\left(\frac{1}{3}\right)^4 \left(\frac{2}{3}\right)^6$.

The question in the previous example is called a series of **Bernoulli trials**. If we have an experiment with only two outcomes, **success**, with probability p, and **failure**, with probability $1 - p$. It is an important assumption in Bernoulli trials that the different trials are independent, i.e. no matter how many times the experiment is repeated, the probabilities of succes and failure don't change.

If we repeat the experiment n times and consider the case of having exactly k successes and $n - k$ failures, then the probability of this is

$$\binom{n}{k} p^n (p-k)^{n-k}.$$

Example 4.2 (Discussion 10.13 in [KT17]). *If we toss a fair coin, the probability of getting heads is $\frac{1}{2}$ (the same with tails). If we toss the coin 100 times, what is the probability that we get exactly 50 heads and 50 tails?*

This is a series of Bernoulli trials with $p = \frac{1}{2}$, so the probability of getting exactly 50 heads and 50 tails is

$$\binom{100}{50} \left(\frac{1}{2}\right)^{50} \left(\frac{1}{2}\right)^{50} = \frac{\binom{100}{40}}{2^{100}} \sim 0.0795892373871788$$

so the probability is approximately $\frac{8}{100}$, not very likely!

5 Discrete Random Variables (Ch. 10.4 in [KT17])

Exercise 5.1. *Consider the situation when we roll two dice and add up the result. If we roll a 2,3,11, or 12, we win \$2. If we roll a 6,7, or 8, we have to pay \$1. The person offering to play argues like this:*

"You win \$2 with 4 out of the 10 outcomes and lose \$1 with only 3 of the possible outcomes. In the long run, you will win!"

Should we play this game?

Solution: Let us compute the probability that we roll a total of 6, 7 or 8 and the probability that we roll a 2,3,11 or 12.

$$\begin{aligned} P(\{6,7,8\}) &= \frac{5}{36} + \frac{6}{36} + \frac{5}{36} \\ &= \frac{16}{36} \\ P(\{2,3,11,12\}) &= \frac{1}{36} + \frac{2}{36} + \frac{2}{36} + \frac{1}{36} \\ &= \frac{6}{36} \end{aligned}$$

So with probability $\frac{16}{36}$ we lose a dollar, and with probability $\frac{6}{36}$ we win two. In the remaining cases (with probability $\frac{14}{36}$), nothing happens. This suggests that we should not play this game.

Definition 5.2. *Let (S, P) be a probability space and let $X : S \to \mathbb{R}$ be any function that maps the outcomes in S to real numbers. We call X a **random variable**. The quantity*

$$E(X) = \sum_{x \in S} X(x)P(x)$$

*is called the **expectation** or **expected value** of the random variable X.*

As the definition suggests, the expectation is the result what one should expect if one performs the experiment. The random variable of the dice game is $X : S \to \mathbb{R}$, given by

$$X(x) = \begin{cases} -1 & \text{if } x \in \{6,7,8\} \\ 0 & \text{if } x \in \{4,5,9,10\} \\ 2 & \text{if } x \in \{2,3,11,12\} \end{cases}$$

and the expectation of the dice game is

$$E(X) = (-1)\left(\frac{16}{36}\right) + (0)\left(\frac{14}{36}\right) + (2)\left(\frac{6}{36}\right) = -\frac{4}{36}.$$

So we can expect to lose $\$\frac{1}{9}$ every time we play.

The moral of the story is: if someone is offering you to play a game where you can win money, then they have probably done the math and you should expect to lose.

Proposition 5.3 (Proposition 10.16. in [KT17])**.** *Let (S, P) be a probability space and let $X_1, X_2, \ldots, X_n$ be random variables. Then*

$$E(X_1 + X_2 + \ldots + X_n) = E(X_1) + E(X_2) + \ldots + E(X_n).$$

Proof. The definition of expectation implies that

$$
\begin{aligned}
E(X_1 + X_2 + \ldots + X_n) &= \sum_{s \in S} (X_1 + X_2 + \ldots + X_n)(x) \\
&= \sum_{s \in S} \sum_{k=1}^{n} X_k(x) \\
&= \sum_{k=1}^{n} \sum_{s \in S} X_k(x) \\
&= \sum_{k=1}^{n} E(X_k)
\end{aligned}
$$

Q.E.D.

6 Central Tendency (Ch. 10.5 in [KT17])

Let's go back to Exercise 5.1. We know that the expectation is that we lose about $\$\frac{1}{9}$ every time we play the game, but how close to the expectation do we expect to be? Let's use a computer simulation.

https://sagecell.sagemath.org/?z=eJxtUE1rwzAMvftXCLNDS9RsTqFdC_kV3W3sYFalMbh2sE2z_vv5i66B-SDZTO_
vSZ60vPv-wHyQLihzuVpD9_6NlbxA2ayMjyVtvad0GawDBcrAp2jbKQl9HRnE46zWvZPmrExYCdytm-dHpqghs3L3Dvf4XlvTKd45bsQDrb
azT4RaFQNH9r9V0DzRvkkLzp88mFycEfqo7R1IY4YXj4hOQI8ghkOOYN0ZeEsyUJgwcy4DIg7qSh7g3zNZwTG4VbOFkE3-UNOoGVbiFj5GA
&lang=sage&interacts=eJyLjgUAARUAuQ==

So it seems that the more we play, the chance of an extreme outcome (winning) should be less.

Definition 6.1. *Let (S, P) be a probability space and X a random variable. The quantity*

$$\operatorname{var}(X) = E\left((X - E(X))^2\right)$$

is the ***variance*** *of X. It is a non-negative number. The quantity*

$$\sigma_X = \sqrt{\operatorname{var}(X)}$$

is the ***standard deviation*** *of X.*

Let's compute the variance and standard deviation for Exercise 5.1. If we play the game just once, the expectation is $\$-\frac{1}{9}$. Let (S, P) and X be as before. The random variable $X - E(X)$ is

$$
(X - E(X))(x) = \begin{cases} -\frac{8}{9} & \text{if } x \in \{6, 7, 8\} \\ \frac{1}{9} & \text{if } x \in \{4, 5, 9, 10\} \\ \frac{19}{9} & \text{if } x \in \{2, 3, 11, 12\} \end{cases}
$$

therefore the random variable $(X - E(X))^2$ is

$$
(X - E(X))^2(x) = \begin{cases} \frac{64}{81} & \text{if } x \in \{6, 7, 8\} \\ \frac{1}{81} & \text{if } x \in \{4, 5, 9, 10\} \\ \frac{361}{81} & \text{if } x \in \{2, 3, 11, 12\} \end{cases}
$$

and therefore the expectation $E\left((X - E(X))^2\right)$ is

$$\operatorname{var}(X) = E\left((X - E(X))^2\right) = \left(\frac{64}{81}\right)\left(\frac{16}{36}\right) + \left(\frac{1}{81}\right)\left(\frac{14}{36}\right) + \left(\frac{361}{81}\right)\left(\frac{6}{36}\right) = \frac{89}{81}$$

therefore the standard deviation is

$$\sigma_X = \sqrt{\operatorname{var}(X)} = \sqrt{\frac{89}{81}}.$$

What are these quantities, the variance and standard deviation measuring? The idea behind standard deviation is that it is supposed to measure how much we expect the result of the experiment to differ from the expectation (this might seem circular at first, but it's a well-defined concept). Let's compute how the variance and standard deviation change if we play the game n times. To do this, we first state some results from probability theory (proofs can be found in Chapter 10.5.1 of [KT17]) that can be used to compute variance.

Proposition 6.2 (Proposition 10.22)**.** *Let X be a random variable in a probability space (S, P). Then* $\text{var}(X) = E(X^2) - (E(X))^2$.

Definition 6.3. *Let X and Y be two random variables in a probability space (S, P). We say that they are* ***independent*** *if for each pair a, b of real numbers with $0 \le a, b \le 1$ the events $\{x \in S | X(x) \le a\}$ and $\{x \in S | Y \le b\}$ are independent.*

Proposition 6.4. *If X and Y are independent random variables, then*

$$\text{var}(X + Y) = \text{var}(X) + \text{var}(Y).$$

For example, if we play the game in Exercise 5.1 n times, then using the above results and the linearity of expectation, we see that (if X is the random variable that is counting how much money we win or lose)

$$E(X) = -\frac{n}{9} \qquad \text{var}(X) = n\left(\frac{89}{81}\right) \qquad \sigma_X = \sqrt{n}\sqrt{(\frac{89}{81}}.$$

So we notice that if we play the game more, the standard deviation also increases, but *not proportionally* to the number of times we play. It increases proportionally to $\sqrt{n}$.

Theorem 6.5 (Chebyshev's inequality, Theorem 10.23. in [KT17])**.** *Let X be a random variable in a probability space (S, P), and let $k \ge 0$ be a positive real number. Then*

$$\text{prob}\left(|X - E(X)| \le k\sigma_X\right) \ge 1 - \frac{1}{k^2}$$

Proof. Let A be the set $\{r \in \mathbb{R} | |r - E(X)| > k\sigma_X\}$.

$$\begin{aligned} \text{var}(X) &= E((X - E(X))^2) \\ &= \sum_{r \in \mathbb{R}} (r - E(X))^2 \text{prob}(X = r) \\ &\ge \sum_{r \in A} (r - E(X))^2 \text{prob}(X = r) \\ &\ge k^2 \sigma_X^2 \sum_{r \in A} \text{prob}(X = r) \\ &\ge k^2 \sigma_X^2 \text{prob}(|X - E(X)| > k\sigma_X). \end{aligned}$$

Since $\text{var}(X) = \sigma_X^2$, we have

$$\frac{1}{k^2} \ge \text{prob}(X - E(X)| > k\sigma_X)$$

which is equivalent to

$$\text{prob}(X - E(X)| \le k\sigma_X) \ge 1 - \frac{1}{k^2}$$

Q.E.D.

What does this Theorem say? The chances of being within k standard deviations of the expectation is at least $1 - \frac{1}{k^2}$. Note that this is only going to be useful for numbers $k > 1$.

Let's apply this to exercise 5.1. If we play the game once, we can just use the probabilities directly to say that we expect to win with probability $\frac{6}{36}$. It's not very good chances, but you would not be shocked to see someone rolling a 2 with two dice.

What are the chances of winning if we play 900 times? We let $n = 900$ in our computation to find that

$$\sigma_X = \sqrt{\frac{80100}{81}} \sim 31.447$$
$$E(X) = -100$$

How can we use Theorem 6.5 to estimate our chances of winning? We win if we end up slightly more than three standard deviations above the expectation. Theorem 6.5 says that we will be within 3 standard deviations at least $\frac{8}{9}$ of the time (note that this also excludes us losing more than \$200), so our chances of winning are now bounded above by $\frac{1}{9}$.

What if we play the game 90000 times? We let $n = 90000$ in our computation to find that

$$\sigma_X = \sqrt{\frac{8010000}{81}} \sim 314.47$$
$$E(X) = -10000$$

Now we need to be more than 31 standard deviations above the expectation to win. Theorem 6.5 now puts the upper bound of our chances of winning at

$$\frac{1}{31^2} = \frac{1}{961}.$$

7 Probability Spaces with Infinitely Many Outcomes (Ch. 10.6 in [KT17])

Exercise 7.1 (Example 10.26 in [KT17])**.** *Consider the following game. You roll a single die, you win if you roll a six, otherwise, the game continues and you roll again until one of the following two situations occurs (1) you roll a six, which now results in a loss or (2) you roll the same number as you got in the first roll, which results in a win. As an example, here are some sequences of rolls that this game might take:*

1. $(4, 2, 3, 5, 1, 1, 1, 4)$ *you win!*
2. (6) *you win!*
3. $(5, 2, 3, 2, 1, 6)$ *you lose.*

What is the probability that you win the game?

Solution: This is not quite a situation to which the machinery we developed applies. The set S of outcomes is not a finite set (the game could go on for an arbitrary long time). But we are still able to assign probabilities to all the different outcomes. For example, the win resulting from (6) has a probability $\frac{1}{6}$ of happening. The win $(4, 2, 3, 5, 1, 1, 1, 4)$ has probability $\frac{1}{6^8}$. Since S is a countably infinite set, we now require P to have the same properties, with the difference that $\sum_{s \in S} P(x)$ is now an infinite sum that should converge to 1. In general, our computation might involve computing sums of infinite series.

We already know that the probability that we win on the first roll is $\frac{1}{6}$. The probability that we win on *exactly* the second roll is

$$\frac{5}{6}\frac{1}{6}$$

since our first roll can't be a six, and the second one has to match the third one. What about winning in exactly three rolls? We need to roll a non-six (this has probability $\frac{5}{6}$), then something other than the first roll or six (probability $\frac{4}{6}$), then we need to roll our first roll (probability $\frac{1}{6}$) for a total probability of

$$\frac{5}{6}\frac{4}{6}\frac{1}{6}.$$

Now it's easy to generalize this to the statement that the probability of winning in exactly n rolls with $n \geq 2$ is

$$\frac{5}{6}\left(\frac{4}{6}\right)^{n-2}\frac{1}{6}.$$

So the probability of winning is

$$\begin{aligned}\frac{1}{6} + \sum_{n \geq 2} \frac{5}{6} \left(\frac{4}{6}\right)^{n-2} \frac{1}{6} &= \frac{1}{6} + \frac{5}{36} \left(\frac{1}{1 - \frac{4}{6}}\right) \\ &= \frac{1}{6} + \frac{5}{36} \frac{1}{\frac{2}{6}} \\ &= \frac{1}{6} + \frac{5}{12} \\ &= \frac{7}{12}\end{aligned}$$

References

[KT17] Mitchel T. Keller and William T. Trotter. *Applied Combinatorics*. Open access, 2017. Available at http://www.rellek.net/appcomb/. 1, 2, 3, 4, 5