# 3. The Geometric Distribution

## Basic Theory

Suppose again that our random experiment is to perform a sequence of Bernoulli trials $\boldsymbol{X} = (X_1, X_2, \ldots)$ with parameter $p \in (0, 1]$. In this section we will study the random variable $U$ that gives the **trial number of the first success**:

$$U = \min\{n \in \mathbb{N}_+ : X_n = 1\}$$

### The Probability Density Function

1. Show that $\{U = n\} = \{X_1 = 0, \ldots, X_{n-1} = 0, X_n = 1\}$

2. Use the result of Exercise 1 and independence to show that the probability density function of $U$ is given by

$$\mathbb{P}(U = n) = p\,(1 - p)^{n-1}, \quad n \in \mathbb{N}_+$$

.

The distribution defined by the probability density in Exercise 2 is known as the **geometric distribution** on $\mathbb{N}_+$, with parameter $p$. The random variable $V = U - 1$ is the number of failures before the first success.

3. Show that the probability density function of $V$ is given by

$$\mathbb{P}(V = n) = p\,(1 - p)^n, \quad n \in \mathbb{N}$$

.

The distribution of this random variable is known as the **geometric distribution** on $\mathbb{N}$, with parameter $p$. Clearly $U$ and $V$ give essentially the same information.

4. In the negative binomial experiment, set $k = 1$ to get the geometric distribution on $\mathbb{N}_+$. Vary $p$ with the scroll bar and note the shape and location of the density function. For selected values of $p$, run the simulation 1000 times with an update frequency of 10. Watch the apparent convergence of the relative frequency function to the density function.

5. Show directly that the function in Exercise 2 is a valid probability density function.

### Moments

The following exercises give the mean, variance, and probability generating function of a random variable $U$ with the geometric distribution on $\mathbb{N}_+$.

6. Show that $\mathbb{E}(U) = \frac{1}{p}$. There are several ways to do this:

a. Derive the result directly using the probability density function.
b. Condition on the first trial to show that $\mathbb{E}(U) = 1 + (1 - p)\,\mathbb{E}(U)$.

7. Show that $\text{var}(U) = \frac{1-p}{p^2}$.

8. Show that $\mathbb{E}\left(t^U\right) = \frac{p\,t}{1-(1-p)\,t}, \quad |t| < \frac{1}{1-p}$.

9. In the negative binomial experiment, set $k = 1$ to get the geometric distribution. Vary $p$ with the scroll bar and note the location and size of the mean/standard deviation bar. For selected values of $p$, run the simulation 1000 times with an update frequency of 10. Watch the apparent convergence of the sample mean and standard deviation to the distribution mean and standard deviation.

10. Suppose now that $V = U - 1$ has the geometric distribution on $\mathbb{N}$. Show that

a. $\mathbb{E}(V) = \frac{1-p}{p}$.
b. $\text{var}(V) = \frac{1-p}{p^2}$.
c. $\mathbb{E}\left(t^V\right) = \frac{p}{1-(1-p)\,t}, \quad |t| < \frac{1}{1-p}$.

### Relation to the Uniform Distribution

Recall that $Y_n$, the number of successes in the first $n$ trials, has the binomial distribution with parameters $n$ and $p$.

11. Show that the conditional distribution of $U$ given $Y_n = 1$ is uniform on $\{1, 2, ..., n\}$. Note that the distribution does not depend on $p$. Interpret the result probabilistically.

### The Memoryless Property

The following problems explore a very important characterization of the geometric distribution.

12. Suppose that $U$ is a random variable taking values in $\mathbb{N}_+$. Show that $U$ has the geometric distribution with parameter $p$ if and only if

$$\mathbb{P}(U > n) = (1 - p)^n, \quad n \in \mathbb{N}$$

13. If $U$ has a geometric distribution on $\mathbb{N}_+$, show that $U$ satisfies the **memoryless property**:

$$\mathbb{P}(U > n + m \mid U > m) = \mathbb{P}(U > n), \quad (n, m) \in \mathbb{N}^2$$

14. Conversely, show that if $U$ is a random variable taking values in $\mathbb{N}_+$ that satisfies the memoryless property, then $U$ has a geometric distribution.

15. Show that $U$ has the memoryless property if and only if the conditional distribution of $U - m$ given $U > m$ is the same as the distribution of $U$.

## Examples and Applications

16. A standard, fair die is thrown until an ace occurs. Let $U$ denote the number of throws.

a. Find the probability density function of $U$
b. Find the mean of $U$.
c. Find the variance of $U$.
d. Find the probability that the die will have to be thrown at least 5 times.

17. A type of missile has failure probability 0.02. Let $N$ denote the number of launches before the first failure.

a. Find the probability density function of $N$
b. Find the mean of $N$.
c. Find the variance of $N$.
d. Find the probability of 20 consecutive successful launches.

18. A student takes a multiple choice test with 10 questions, each with 5 choices (only one correct). The student blindly guesses and gets one question correct. Find the probability that the correct question was one of the first 4.

19. Recall that an American roulette wheel has 38 slots: 18 are red, 18 are black, and 2 are green. Suppose that you observe red or green on 10 consecutive spins. Give the conditional distribution of the number of additional spins needed for black to occur.

The game of roulette is studied in more detail in the chapter on Games of Chance.

20. In the negative binomial experiment, set $k = 1$ to get the geometric distribution and set $p = 0.3$. Run the experiment 1000 times, with an update frequency of 100. Compute the appropriate relative frequencies and empirically investigate the memoryless property

$$\mathbb{P}(U > 5 | U > 2) = \mathbb{P}(U > 3)$$

### The Petersburg Problem

We will now explore a gambling situation, known as the **Petersburg problem**, which leads to some famous and surprising results. Suppose that we are betting on a sequence of Bernoulli trials with success parameter

$p > 0$. We can bet any amount of money on a trial at **even stakes**: if the trial results in success, we receive that amount, and if the trial results in failure, we must pay that amount. We will use the following strategy, known as a **martingale strategy**:

1. We bet $c$ monetary units on the first trial.
2. Whenever we lose a trial, we double the bet for the next trial.
3. We stop as soon as we win a trial.

21. Let $W$ denote our net winnings when we stop. Show that $W = c$.

Thus, $W$ is not random and $W$ is independent of $p$! Since $c$ is an arbitrary constant, it would appear that we have an ideal strategy. However, let us study the amount of money $Z$ needed to play the strategy.

22. Show that $Z = c\left(2^U - 1\right)$. where as usual, $U$ is the trial number of the first success.

23. Use the result in the previous exercise to show that

$$\mathbb{E}(Z) = \begin{cases} \dfrac{c}{2p - 1}, & p > \dfrac{1}{2} \\ \infty, & p \le \dfrac{1}{2} \end{cases}$$

Thus, the strategy is fatally flawed when the trials are unfavorable and even when they are fair.

24. Compute $\mathbb{E}(Z)$ explicitly if $c = 100$ and $p = 0.55$.

25. In the negative binomial experiment, set $k = 1$. For each of the following values of $p$, run the experiment 100 times, updating after each run. For each run compute $Z$ (with $c = 1$). Find the average value of $Z$ over the 100 runs:

a. $p = 0.2$
b. $p = 0.5$
c. $p = 0.8$

For more information about gambling strategies, see the chapter on Red and Black.

### The Alternating Coin-Tossing Game

A coin has probability of heads $p \in (0, 1]$. There are $n$ players who take turns tossing the coin in round-robin style: player 1 first, then player 2, ..., then player $n$, then player 1 again, and so forth. The first player to toss heads wins the game.

Let $U$ denote the number of the first toss that results in heads. Of course, $U$ has the geometric distribution on

$\mathbb{N}_+$ with parameter $p$. Additionally, let $W$ denote the winner of the game; $W$ takes values in the set $\{1, 2, ..., n\}$. We will compute the probability density function of $W$ in two different ways

26. Show that for $i \in \{1, 2, ..., n\}$, $W = i$ if and only if $U = i + k\,n$ for some $k \in \mathbb{N}$. That is, using modular arithmetic,

$$W = ((U - 1) \bmod n) + 1$$

.

27. Use the result of the previous exercise and the geometric distribution to show that

$$\mathbb{P}(W = i) = \frac{p\,(1 - p)^{i-1}}{1 - (1 - p)^n}, \quad i \in \{1, 2, ..., n\}$$

28. Argue that $\mathbb{P}(W = i) = (1 - p)^{i-1}\,\mathbb{P}(W = 1)$. Use this result to re-derive the probability density function in the previous exercise.

29. Explicitly compute the probability density function of $W$ when the coin is fair ($p = \frac{1}{2}$)

Note from Exercise 27 that $W$ itself has a **truncated geometric distribution**.

30. Show that the distribution of $W$ is the same as the conditional distribution of $U$ given $U \le n$:

$$\mathbb{P}(W = i) = \mathbb{P}(U = i | U \le n), \quad i \in \{1, 2, ..., n\}$$

The following problems explore some limiting distributions related to the alternating coin-tossing game.

31. Show that for fixed $p \in (0, 1]$, the distribution of $W$ converges to the geometric distribution with parameter $p$ as $n \uparrow \infty$.

32. Show that for fixed $n$, the distribution of $W$ converges to the uniform distribution on $\{1, 2, ..., n\}$ as $p \downarrow 0$.

33. What happens in the game when $p = 0$? Compare with the limit in the previous exercise.

---