

Quiz II. MEMORYLESS DISTRIBUTIONS

A random variable X is said to have an *exponential distribution* with parameter δ iff it has pdf

$$f(x) = \begin{cases} \delta e^{-\delta x}, & x > 0 \\ 0, & x \leq 0. \end{cases}$$

Thus, an exponential distribution is just a gamma distribution with parameters

$$\alpha = \underline{\hspace{2cm}}, \quad \beta = \underline{\hspace{2cm}}.$$

Accordingly, such an X has mean and variance

$$\mu = \underline{\hspace{2cm}}, \quad \sigma^2 = \underline{\hspace{2cm}},$$

and m.g.f.

$$M(t) = \underline{\hspace{2cm}} \text{ for } t \underline{\hspace{2cm}}.$$

We will see that such distribution has no memory. To begin, note that if X has an exponential distribution with parameter δ , then for every number $t > 0$,

$$P(X \geq t) = \underline{\hspace{2cm}}.$$

(Note: do *not* confuse this with the d.f.)

Therefore,

$$Pr(X \geq t + h | X \geq t) = Pr(X \geq h), \tag{1}$$

showing that $Pr(X \geq t + h | X \geq t)$ is independent of t .

Prove that $Pr(X \geq t + h | X \geq t) = Pr(X \geq h)$ in the space below. (If you need inspiration, see #2.18 or #3.61.)

To illustrate the memoryless property, suppose that X represents the number of minutes that elapse before some event occurs. According to (2), if the event has not occurred in t minutes, then the probability that the event will not occur during the next h minutes is¹

$$Pr(X \geq t + h | X \geq t) = \underline{\hspace{2cm}}.$$

This is just the same as the probability that the event would not occur during an interval of h minutes, starting at time 0! In other words, regardless of the time that has elapsed, the probability that the event will happen in the next h minutes is always the same. (Think about this for a moment — it is kinda weird.) Theoretically, it is therefore not necessary to consider past occurrences of an event, in order to calculate probabilities for future occurrences.

¹Give me the formula you found during the above proof, i.e., do *not* write $Pr(X \geq h)$.

Life Tests

Suppose n light bulbs are burn simultaneously in a test to determine the lengths of their lives. Let X_i denote the lifetime of the i^{th} bulb. Then X_1, \dots, X_n are i.i.d. with exponential distribution of parameter δ .

Determine the distribution of the length of time Y_1 until one of the n bulbs fails.

$$Pr(Y_1 > t) = \underline{\hspace{2cm}}.$$

This is what kind of a distribution? _____

With what parameter values? _____

After one bulb has failed, $n - 1$ bulbs are still burning. Furthermore, regardless of what time the first failed, the memoryless property indicates that the distribution of the remaining lifetime of each of the remaining $n - 1$ bulbs is still exponential with parameter δ . In other words, the situation is the same as if we started the experiment over again at time $t = 0$ with $n - 1$ new bulbs.

Let Y_2 be the interval of time between the failure of the first bulb and the second.

What is the distribution of Y_2 ? _____

After all but one of the bulbs have failed what is the distribution of the interval until the final bulb dies? _____

Radioactive Particles

Suppose we are interested in how long we have to wait until a radioactive particle strikes a target. Let Y be time until the first particle strikes the target, and let X be the number of particles that strike the target in time t . Then

$$Y \leq t \quad \Longleftrightarrow \quad X \geq 1.$$

Suppose X has a Poisson distribution with mean λt , where λ is the rate of the process. Find the df of Y (If you need inspiration, see Example 1 on page 129.)

$$F(t) = Pr(Y \leq t) = \begin{cases} \text{_____}, & \text{for } t \text{ _____}, \\ \text{_____}, & \text{for } t \text{ _____}. \end{cases}$$

Taking the derivative, we find the pdf

$$f(t) = \begin{cases} \underline{\hspace{2cm}}, & \text{for } t \underline{\hspace{1cm}}, \\ \underline{\hspace{2cm}}, & \text{for } t \underline{\hspace{1cm}}. \end{cases}$$

But this is just the pdf of a _____ distribution with parameter values _____!

Uniqueness

Now we will prove that any continuous distribution which is memoryless must be an exponential distribution. Please write up your proofs on separate paper and staple it to your quiz when you turn it in. Remember: these problems are given in order so that you can use earlier parts in solving later parts.

Let F be a continuous distribution function satisfying $F(0) = 0$ and suppose that the distribution with df F has the memoryless property:

$$Pr(X \geq t + h | X \geq t) = Pr(X \geq h).$$

Define a new function

$$\ell(x) = \log[1 - F(x)].$$

1. Show that for any $t, h > 0$,

$$1 - F(h) = \frac{1 - F(t + h)}{1 - F(t)}.$$

2. Use this to show that for any $t, h > 0$,

$$\ell(t + h) = \ell(t) + \ell(h).$$

3. Use this to show that for any $t > 0$ and all positive *integers* k and m ,

$$\ell\left(\frac{kt}{m}\right) = \frac{k}{m}\ell(t).$$

Hint: first show $\ell(kt) = k\ell(t)$ for any $t > 0$. Then, use this fact and the fact that $1 = m \cdot \frac{1}{m}$ to show $\ell\left(\frac{1}{m}t\right) = \frac{1}{m}\ell(t)$. Then put them together.

4. Use this to show that for any $t, c > 0$,

$$\ell(ct) = c\ell(t).$$

Hint 1: since c can be any positive real number, it is helpful to know that we can always find a sequence of rational numbers $\{r_n\}_{n=1}^{\infty}$ which converge to c . In other words, we have numbers of the form $r_n = \frac{k_n}{m_n}$ for $n = 1, 2, \dots$ where each k_n and m_n are positive integers; and $\lim_{n \rightarrow \infty} r_n = c$.

Hint 2: a function $h(x)$ is continuous iff

$$\lim_{n \rightarrow \infty} h(x_n) = h(c), \text{ for any sequence } \{x_n\}_{n=1}^{\infty} \text{ with } \lim_{n \rightarrow \infty} x_n = c.$$

5. Prove that $g(t) = \frac{\ell(t)}{t}$ is constant for $t > 0$. In fact, show that for $t > 0$, $g(t) = -\delta$, where $\delta > 0$.
6. Prove that F must be the df of an exponential distribution.

But wait, there's more!

There are other memoryless distributions, but they are discrete (they would have to be - we just proved it on the last page!). Consider a sequence of Bernoulli trials in which the outcome of each trial is either a success or a failure and the probability of a success on each trial is p , where $0 < p < 1$. Then the distribution of the *number of failures which will occur before the first success* is a geometric distribution with parameter p . This distribution is discussed on page 121.

Suppose now that a failure occurred on each of the first 20 trials. Then, since all trials are independent, the distribution of the *additional* failures which will occur before the first success is obtained will again be a geometric distribution with parameter p . In effect, the process begins anew with the 21st trial, and the long sequence of failures that were obtained on the first 20 trials have no effect on the future outcomes of the process.

Thus the geometric distribution is memoryless, as we will show. It also has the pdf

$$f(x) = p(1 - p)^x, \quad x = 0, 1, 2, \dots$$

7. Show that for an rv X with geometric distribution,

$$Pr(X \geq k) = (1 - p)^k.$$

Hint: recall that for $0 < r < 1$,

$$\sum_{n=0}^{k-1} r^n = \frac{1 - r^k}{1 - r}, \quad \text{and} \quad \sum_{n=0}^{\infty} r^n = \frac{1}{1 - r}.$$

8. Use this to show that X has no memory, i.e.,

$$Pr(X = t + k | X \geq k) = Pr(X = t).$$

This phenomenon leads to one of the classic quandaries that people, especially gamblers, encounter in probability. The odds of getting a failure *every* time are small, especially for a large number of trials. Therefore, people believe that if they play a game long enough, they are bound to win eventually. Intuitively, the longer a gambler has been losing at the tables, the more likely they feel they are to win on the next round. However, the existence of memoryless distributions shows this to be a mistaken belief.

This quiz is due in discussion on February 10. Feel free to ask me about it, but please do not discuss with other students. Good luck!