# Practice Problems in Probability

## Easy and Medium Difficulty Problems

**Problem 1.** Suppose we flip a fair coin once and observe either T for "tails" or H for "heads." Let $X_1$ denote the random variable that equals 0 when we observe tails and equals 1 when we observe heads. (This is called a *Bernoulli* random variable.)

(a) Make a table of the PDF of $X_1$ and calculate $\mathbb{E}(X_1)$ and $\mathrm{Var}(X_1)$.

(b) Now suppose that we flip a fair coin twice. We will observe one of four events: TT, TH, HT, or HH. Let $X_2$ be the random variable that counts the number of heads we observe. (So $X_2$ can take the values 0, 1, or 2.) Redo part (a) for this new random variable.

(c) Let $X_3$ be the random variable that counts the number of heads we observe after three successive flips of a fair coin. Redo part (a) for this new random variable.

(d) Do you notice any patterns in your calculations? Make a guess for the pdf table, the expectation and the variance of $X_4$, the random variable that counts the number of heads we observe after four successive flips of a fair coin, and then verify your guess by direct computation.

**Problem 2.** For any positive integer $n$, the random variable $X_n$ defined in Problem 1 is called a *binomial* random variable. The PDF of the random variable is given by

$$\Pr(X_n = k) = \frac{n!}{k!(n-k)!}\left(\frac{1}{2}\right)^n,$$

where $k$ is the number of heads (or tails) in $n$ successive and independent flips of a fair coin. (Recall that the *factorial* notation denotes a product of integers: $n! = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n$.)

(a) Calculate $\Pr(X_{10} = 5)$. Interpret this probability.

(b) Calculate $\Pr(X_{100} = 1)$ and $\Pr(X_{100} = 99)$. Interpret this result.

(c) As we saw in the learning module, we can think of the flip of a coin as an experimental trial and the outcome - heads or tails - as a "success" or "failure". Generally speaking then, we can imagine observing the outcomes of $n$ independent trials of some repeatable experiment, each time observing a "successful" outcome with probability $p$. The number of successful trials $X_n$ is thus given by a generic binomial random variable with PDF

$$\Pr(X_n = k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k},$$

where $k$ is the number of successes observed in $n$ independent trials. Calculate $\mathbb{E}(X_n)$, the expected number of successes in $n$ independent trials. Does this agree with your intuition?

**Problem 3.** Consider a slightly different coin tossing experiment. Suppose we toss a fair coin and continue to toss it until we first observe "heads". If we let $Y$ denote the random variable that counts the number of tosses until we first observe heads, then we see the possible values of $Y$ are $1,\ 2,\ 3, \ldots$. (This is an example of a *geometric* random variable.)

(a) Construct a PDF table for $Y$. Write down a formula for $\Pr(Y = y)$ for any positive integer $y$.

(b) Construct a CDF table for $Y$. Use this to deduce that

$$\sum_{k=1}^{\infty} \left(\frac{1}{2}\right)^k = 1.$$

**Problem 4.** In a manufacturing process, suppose that the probability that we produce a defective item is $\frac{1}{100}$. Let us observe the items on the production line until we find the first defective item. Denote this random variable by $X$ and observe that $X$ is a geometric random variable as in Problem 4.

(a) Construct a PDF table for $X$ and compute $\Pr(X \leq 5)$. What is $\Pr(X = k)$ for any positive integer $k$?

(b) Write down the CDF for $X$; that is, write down a formula for $\Pr(X \leq k)$ for any positive integer $k$. Use this to show explicitly that $\lim_{k\to\infty} \Pr(X \leq k) = 1$.

(c) Compute $\Pr(1 < X \leq 10)$ and $\Pr(X > 10)$.

**Problem 5.** Recall the example of rolling a six-sided die. This is an example of a *discrete uniform* random variable, so named because the probability of observing each distinct outcome is the same, or uniform, for all outcomes. Let $Y$ be the discrete uniform random variable that equals the face-value after a roll of an *eight*-sided die. (The die has eight faces, each with a number 1 through 8.) Calculate $\mathbb{E}(Y)$, $\mathrm{Var}(Y)$, and $\mathrm{StdDev}(Y)$.

**Problem 6.** Given the following table:

| $x$ | $\Pr(X \le x)$ |
|---|---|
| $(-\infty, 1)$ | 0 |
| $[1, 3)$ | 1/4 |
| $[3, 4)$ | 3/7 |
| $[4, 7)$ | 2/3 |
| $[7, 8)$ | 7/8 |
| $[8, \infty)$ | 1 |

(a) Graph the CDF and PDF of $X$.

(b) Find a constant $c$ such that $\Pr(X \le c) = \frac{7}{8}$.

(c) Find a constant $c$ such that $\Pr(X < c) = \frac{2}{3}$.

(d) Find a constant $c$ such that $\Pr(X > c) = \frac{4}{7}$.

(e) Find constants $c_1$, $c_2$ such that $\Pr(c_1 \le X < c_2) = \frac{3}{4}$.

**Problem 7.** Using properties of sums, show that $\mathrm{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$, for any discrete random variable $X$.

**Problem 8.** Repeat Problem 7 for any continuous random variable $X$, using properties of integrals.

**Problem 9.** Are the following functions PDF's?

(a) $f(x) = \begin{cases} 12x^2(x-1) & \text{if } 0 < x < 1 \\ 0 & \text{if otherwise} \end{cases}$

(b) $f(x) = \begin{cases} 1 - |x| & \text{if } |x| \le 1 \\ 0 & \text{if otherwise} \end{cases}$

(c) $f(x) = \begin{cases} \frac{\pi}{2}\cos(\pi x) & \text{if } |x| < \frac{1}{2} \\ 0 & \text{if otherwise} \end{cases}$

(d) $f(x) = \begin{cases} \frac{1}{2} & \text{if } |x| \le 2 \\ 0 & \text{if otherwise} \end{cases}$

**Problem 10.** For the functions in Problem 9 that you found to be probability density functions, find the corresponding cumulative distribution functions.

**Problem 11.** Are the following functions CDF's?

(a) $F(x) = \begin{cases} 0 & \text{if } x < -\frac{\pi}{2} \\ \frac{\sin(x-\pi/2)+1}{2} & \text{if } |x| \leq \frac{\pi}{2} \\ 1 & \text{if } x > \frac{\pi}{2} \end{cases}$

(b) $F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1-\cos(x)}{2} & \text{if } 0 \leq x \leq \pi \\ 1 & \text{if } x > \pi \end{cases}$

(c) $F(x) = \begin{cases} 0 & \text{if } x < -1 \\ 1+x & \text{if } -1 \leq x < 0 \\ x & \text{if } 0 \leq x \leq 1 \\ 1 & \text{if } x > 1 \end{cases}$

(d) $F(x) = \arctan(x) + \frac{\pi}{2}$

**Problem 12.** For the functions in Problem 11 that you found to be cumulative distribution functions, find the corresponding probability density functions.

**Problem 13.** Show that $e^{-x}e^{-e^{-x}}$ on $x \in \mathbb{R}$ is a PDF.

**Problem 14.** What is the CDF of the density function $\frac{1}{\pi(1+x^2)}$?

**Problem 15.** Show that $p(x) = \frac{e^{-x}}{(1+e^{-x})^2}$ on $x \in \mathbb{R}$ is a PDF.

**Problem 16.** Show that $f(x) = \begin{cases} 1 - e^{-x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$ is a CDF.

**Problem 17.** Find the constant $k$ that makes the following functions PDF's on the given domains.

(a) $p(x) = k\sin(x)$, $0 < x < \pi$

(b) $p(x) = kx^2(x-1)^2$, $0 < x < 1$

(c) $p(x) = kx(1-x)^3$, $0 < x < 1$

(d) $p(x) = k$, $-1 \leq x \leq 3$

(e) $p(x) = kx^3e^{-\frac{x}{2}}$, $x \geq 0$.

**Problem 18.** For the PDF's in Problem 17, compute the expectations, variances and standard deviations of their associated random variables.

**Problem 19.** Let $Z$ be a *standard normal* random variable (the classic "bell curve"), given by the density

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Verify that $\mathbb{E}(Z) = 0$.

**Problem 20.** Expectations (and variances) need not be finite. Let $Y$ be a *Cauchy* random variable given by the PDF: $p(x) = \frac{1}{\pi(1+x^2)}$, for $x \in \mathbb{R}$.

(a) Prove that $\mathbb{E}(X)$ does not exist (i.e. the expectation is infinite).

(b) Prove that $\mathbb{E}(X^2)$ does not exist.

**Problem 21.** Let $p(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}}$, for $0 \leq x < \infty$, $\beta > 0$, an *exponential* random variable with parameter $\beta$.

(a) Show that this density integrates to 1.

(b) Calculate $\Pr(X \geq 1)$, $\mathbb{E}(X)$ and $\mathrm{StdDev}(X)$, for the exponential random variable $X$.

(c) Sketch the PDF of $X$ for $\beta = \frac{1}{10},\ \frac{1}{2},\ 1,\ 5.$

(d) Find the CDF of $X$ in general.

**Problem 22.** Show that, for any continuous random variable $X$, $\Pr(X \leq x) = \Pr(x < x)$. Give a counterexample to show that this equality does not hold if $X$ is a discrete random variable.

**Problem 23.** Prove that $\Pr(X \leq x) = 1 - \Pr(X > x)$ first for discrete random variables, then for continuous random variables.

**Problem 24.** Let $X$ be a *Laplace* random variable given by the PDF: $p(x) = \frac{1}{2} e^{-|x|}$, for $x \in \mathbb{R}$.

(a) Verify that $p(x)$ is in fact a valid pdf.

(b) Calculate $\Pr(|X| \geq 1)$, $\mathbb{E}(X)$, and $\mathrm{StdDev}(X)$.

(c) Sketch the PDF of $X$.

(d) Find an explicit formula for the CDF of $X$.

**Problem 25.** Let $Y$ be a *continuous uniform* random variable on the interval $[a, b]$.

(a) Graph the PDF and CDF of $Y$. Is the CDF continuous everywhere? Is the CDF differentiable everywhere?

(b) Calculate $\mathbb{E}(Y)$, $\mathrm{Var}(Y)$, and $\mathrm{StdDev}(Y)$.

**Problem 26.** Show that $\mathbb{E}(X) \leq \sqrt{\mathbb{E}(X^2)}$ for any random variable $X$.

**Problem 27.** Let $X$ be a normal random variable with mean $\mu = -1$ and variance $\sigma^2 = 100$. Given that $\Pr(Z \geq 1) = 0.1587$, where $Z$ is a standard normal random variable,

(a) calculate $\Pr(X \geq 9)$

(b) calculate $\Pr(-11 \leq X \leq 9)$

(c) calculate $\Pr(|X+1| > 10)$

**Problem 28.** *(from "Mathematical Statistics with Applications", by Wackerly, Mendenhall, and Scheaffer)*
The weekly amount of money spent on maintenance and repairs by a company was observed, over a long period of time, to be approximately normally distributed with mean \$400 and standard deviation \$20. If \$450 is budgeted for next week, what is the probability that the actual costs will exceed the budgeted amount? You may use the fact that $\Pr(Z \leq -2.5) = 0.0062$, where $Z$ is a standard normal random variable.

**Problem 29.** The final numerical grades of college students in a certain large biology class are approximately normally distributed with mean 74.1 and standard deviation 9.6. You may use the fact that $\Pr(Z \geq 1.47) = 0.0708$, where $Z$ is a standard normal random variable.

(a) What percentage of students would we expect to pass the course if a final grade of 60 or higher is required to pass?

(b) What percentage of students would we expect to receive a final grade greater than the mean?

(c) If final grades were *not* normally distributed, but still had a mean of 74.1 and standard deviation of 9.6, would your answer to part (b) change?

**A Few Challenging Problems:**

**Problem 30.** A *Poisson* random variable $X$ is given by the pdf

$$\Pr(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \text{ for } \lambda \geq 0,\ x = 0, 1, 2, \ldots.$$

(a) Use Taylor series to show that $\sum_{x=0}^{\infty} \Pr(X = x) = 1$.

(b) Use Taylor series to show that $\mathbb{E}(X) = \lambda$.

(c) Use Taylor series to show that $\text{Var}(X) = \lambda$. (Hint: Compute $\mathbb{E}[X(X-1)]$.)

**Problem 31.** Technically, a function $f(x)$ is a *cumulative distribution function* if and only if it is nondecreasing, $\lim_{x\to-\infty} f(x) = 0$, $\lim_{x\to\infty} f(x) = 1$ *and* if $f(x)$ is continuous from the right. For example, the function $f_1(x) = \begin{cases} 1 & \text{if } x \geq 0, \\ 0 & \text{if } x < 0 \end{cases}$ is a CDF, while the function $f_2(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x \leq 0 \end{cases}$ is not a CDF. Considering these examples, why does this technical requirement make intuitive sense?

**Problem 32.** Recall that the average value of a function on the interval $[a, b]$ is given by $f_{avg}(a, b) = \frac{1}{b-a} \int_a^b f(x)dx$. Let $X$ be a (continuous) uniform random variable on $[a, b]$. Define the new random variable $Y = f(X)$ for any strictly increasing function $f$. Show that $\mathbb{E}(Y) = f_{avg}(a, b)$. Interpret this result.