MATH 461 LECTURE NOTE
WEEK 4

DAESUNG KIM

1. Other Discrete Random Variables (Sec 4.8)

Geometric random variables. A random variable X is called geometric if it takes values $1, 2, \cdots$ with probabilities

$$\mathbb{P}(X = n) = p(1-p)^{n-1}$$

for some $p \in (0, 1)$ and denoted by $X \sim \mathrm{Geom}(p)$. Note that $\sum_{n=1}^{\infty} p(1-p)^{n-1} = 1$.

Suppose we perform a sequence of independent Bernoulli experiments with success probability $p \in [0, 1]$. Let X be the number of trials until the first success occurs. Then, X is a geometric random variable with p.

Example 1. An urn contains N white and M black balls. Balls are randomly selected, one at a time, until a black one is obtained. If we assume that each ball selected is replaced before the next one is drawn, what is the probability that (a) exactly n draws are needed? (b) at least k draws are needed?

Proposition 2. *Let $X \sim \mathrm{Geom}(p)$ for some $p \in (0, 1)$.*

(i) *For any two nonnegative integers $m, n \geq 0$,*

$$\mathbb{P}(X > m + n | X > m) = \mathbb{P}(X > n).$$

(ii) $\mathbb{E}[X] = \frac{1}{p}$ *and* $\mathrm{Var}(X) = \frac{1-p}{p^2}$.

Proof. (i) It follows from $\mathbb{P}(X > k) = (1-p)^{k+1}$.

(ii) We have

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{k=1}^{\infty} k\mathbb{P}(X = k) \\
&= \sum_{k=1}^{\infty} kp(1-p)^{k-1} \\
&= (1-p)\sum_{k=1}^{\infty}(k-1)p(1-p)^{k-2} + 1 \\
&= (1-p)\mathbb{E}[X] + 1
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}[X^2] &= \sum_{k=1}^{\infty} k^2\mathbb{P}(X = k) \\
&= (1-p)\sum_{k=1}^{\infty}(k-1)^2p(1-p)^{k-2} + 2\mathbb{E}[X] - 1 \\
&= (1-p)\mathbb{E}[X^2] + \frac{2}{p} - 1.
\end{aligned}$$

Thus, $\mathrm{Var}(X) = \frac{1-p}{p^2}$.

□

Negative binomial random variables. Let $p \in (0,1)$ and $r \geq 1$ be an integer. A random variable X is called negative binomial if it takes values $r, r+1, \cdots$ with probabilities

$$\mathbb{P}(X = n) = \binom{n-1}{r-1} p^r (1-p)^{n-r}.$$

Suppose we perform a sequence of independent Bernoulli experiments with success probability $p \in [0,1]$. Let X be the number of trials until success occurs r times. Then, X is a negative binomial random variable with r and p. Note that a geometric random variable is negative binomial with $r = 1$. One can show that $\mathbb{E}[X] = \frac{r}{p}$ and $\mathrm{Var}(X) = \frac{r(1-p)}{p^2}$.

Example 3 (The Banach match problem)**.** At all times, a pipe-smoking mathematician carries 2 matchboxes: 1 in his left-hand pocket and 1 in his right-hand pocket. Each time he needs a match, he is equally likely to take it from either pocket. Consider the moment when the mathematician first discovers that one of his matchboxes is empty. If it is assumed that both matchboxes initially contained N matches, what is the probability that there are exactly k matches, $k = 0, 1, \cdots, N$, in the other box?

Hypergeometric Random variables. Let N, m, n be positive integers with $N > n, m$. A random variable X is called hypergeometric if it takes values $0, 1, 2, \cdots, n$ with probabilities

$$\mathbb{P}(X = i) = \frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}}$$

for $i = 0, 1, 2, \cdots, n$.

Suppose that a sample of size n is to be chosen randomly (without replacement) from an urn containing N balls, of which m are white and $N - m$ are black. Let X denote the number of white balls selected, then X is hypergeometric. If we choose n balls with replacement and Y the number of white balls selected, then Y is binomial with n and m/N.

2. Intro to Continuous Random Variables (Sec 5.1)

A discrete random variable is a random variable whose possible values are countable. However, one can think a random variable with uncountable possible values, for instance, the lifetime of a light bulb. For a discrete random variable, we assign the probabilities for each possible values, which is not possible for the case with uncountable possible values. Then, how can we define the probabilities for this case?

Definition
A random variable X is a continuous random variable if there is a nonnegative function f on $\mathbb{R}$ such that

$$\mathbb{P}(X \in B) = \int_B f(x)\, dx$$

for any set $B \subseteq \mathbb{R}$. The function f is called the probability density function of X.

In particular, we have

$$\begin{aligned}
\mathbb{P}(X \in \mathbb{R}) &= \int_{\mathbb{R}} f(x)\, dx = 1, \\
\mathbb{P}(a \leq X \leq b) &= \int_a^b f(x)\, dx, \\
\mathbb{P}(X = a) &= \int_a^a f(x)\, dx = 0.
\end{aligned}$$

Example 4. Suppose that X is a continuous random variable whose probability density function is given by

$$f(x) = \begin{cases} C(3x - x^2), & 0 < x < 3, \\ 0, & \text{otherwise.} \end{cases}$$

Find C and $\mathbb{P}(X > 1)$.

The cumulative distribution function (in short, cdf) of a random variable X is defined by

$$F(a) = \mathbb{P}(X \le a) = \mathbb{P}(X < a) = \int_{-\infty}^{a} f(x)\,dx.$$

That is, the cdf is the integral of the probability density of X. Note that the cdf is continuous (regardless of the continuity of f). On the other hands, if we differentiate $F(a)$ with respect to a, then

$$\frac{d}{da}F(a) = \frac{d}{da}\int_{-\infty}^{a} f(x)\,dx = f(a).$$

That is, the probability density is the derivative of the cdf. Note that, for small $\varepsilon > 0$, we have

$$\mathbb{P}\left(a - \frac{\varepsilon}{2} \le X \le a + \frac{\varepsilon}{2}\right) = \int_{a-\frac{\varepsilon}{2}}^{a+\frac{\varepsilon}{2}} f(x)\,dx \approx \varepsilon f(a).$$

Example 5. Let X be a continuous random variable with the density f_X and the cdf F_X. Let $Y = X^2$. What is the density of Y?

3. Expectation and Variance of Continuous RVs (Sec 5.2)

> **Expectation**
> Let X be a continuous random variable with the density f. The expectation (or mean, or the expected value) of X is defined by
> $$\mathbb{E}[X] = \int_{\mathbb{R}} x f(x)\,dx.$$

Example 6. Find $E[X]$ when the density function of X is given by

$$f(x) = \begin{cases} 2x, & 0 \le x \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

Example 7. Let X be a random variable with density

$$f(x) = \begin{cases} 2x, & 0 \le x \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

Let $Y = e^X$.

(i) Find the density f_Y and distribution function F_Y of Y.
(ii) Compute $\mathbb{E}[Y]$.

Proposition 8. *Let X be a continuous random variable with the density f.*

(i) *If $X \ge 0$, then*
$$\mathbb{E}[X] = \int_0^{\infty} \mathbb{P}(X > x)\,dx.$$
(ii) *If g a real-valued function, then $\mathbb{E}[g(X)] = \int_{\mathbb{R}} g(x)f(x)\,dx$. In particular $\mathbb{E}[aX+b] = a\mathbb{E}[X] + b$ for $a, b \in \mathbb{R}$.*

Proof. (i)

$$\begin{aligned}\int_0^{\infty} \mathbb{P}(X > x)\,dx &= \int_0^{\infty}\int_x^{\infty} f(y)\,dydx \\ &= \int_0^{\infty}\int_0^{y} f(y)\,dxdy \\ &= \int_0^{\infty} y f(y)\,dy \\ &= \mathbb{E}[X].\end{aligned}$$

(ii) Suppose g is nonnegative. Since $Y = g(X)$ is nonnegative, it follows from (i) that

$$\begin{aligned}\mathbb{E}[g(X)] &= \int_0^\infty \mathbb{P}(g(X) > y)\, dy \\ &= \int_0^\infty \int_{x:g(x)>y} f(x)\, dxdy \\ &= \int \int_0^{g(x)} f(x)\, dydx \\ &= \int g(x) f(x)\, dx.\end{aligned}$$

For general g, one can use

$$\mathbb{E}[Y] = \int_0^\infty \mathbb{P}(Y > y)\, dy - \int_0^\infty \mathbb{P}(Y < -y)\, dy$$

(see [SR, p.215, Problem 5.2]).

□

Example 9 (Revisit)**.** Compute $\mathbb{E}[e^X]$ where X is a random variable with density

$$f(x) = \begin{cases} 2x, & 0 \le x \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

Example 10. Consider an interval $[0, 1]$ and a point $p \in [0, 1]$. Let X be a random variable with density

$$f(x) = \begin{cases} 1, & 0 \le x \le 1, \\ 0, & \text{otherwise} \end{cases}$$

and consider two subintervals $[0, X]$ and $[X, 1]$. Let L be the length of one of these subintervals that contains p.

(i) Find a function g such that $L = g(X)$.
(ii) Compute $\mathbb{E}[L]$.

Variance
The variance is defined by

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \int_{\mathbb{R}} (x - \mathbb{E}[X])^2 f(x)\, dx.$$

References

[SR] Sheldon Ross, *A First Course in Probability*, 9th Edition, Pearson

Department of Mathematics, University of Illinois at Urbana-Champaign
E-mail address: daesungk@illinois.edu