Lecture 27

1. Two important properties

Theorem 27.1 (Uniqueness). *If X and Y are two random variables—discrete* or *continuous—with moment generating functions M_X and M_Y, and if there exists $\delta > 0$ such that $M_X(t) = M_Y(t)$ for all $t \in (-\delta\,, \delta)$, then $M_X = M_Y$ and X and Y have the same distribution. More precisely:*

1. *X is discrete if and only if Y is, in which case their mass functions are the same;*
2. *X is continuous if and only if Y is, in which case their density functions are the same.*

Theorem 27.2 (Lévy's continuity theorem). *Let X_n be a random variables—discrete* or *continuous—with moment generating functions M_n. Also, let X be a random variable with moment generating function M. Suppose there exists $\delta > 0$ such that:*

1. *If $-\delta < t < \delta$, then $M_n(t), M(t) < \infty$ for all $n \geq 1$; and*
2. $\lim_{n\to\infty} M_n(t) = M(t)$ *for all $t \in (-\delta\,, \delta)$, then*

$$\lim_{n\to\infty} F_{X_n}(a) = \lim_{n\to\infty} \mathrm{P}\left\{X_n \leq a\right\} = \mathrm{P}\{X \leq a\} = F_X(a),$$

for all numbers a at which F_X is continuous.

Example 27.3 (Law of rare events). Suppose $X_n = \text{binomal}(n\,, \lambda/n)$, where $\lambda > 0$ is fixed, and $n \geq \lambda$. Then, recall that

$$M_{X_n}(t) = \left(q + pe^{-t}\right)^n = \left(1 - \frac{\lambda}{n} + \frac{\lambda e^{-t}}{n}\right)^n \to \exp\left(-\lambda + \lambda e^{-t}\right).$$

Note that the right-most term is $M_X(t)$, where $X = \text{Poisson}(\lambda)$. Therefore, by Lévy's continuity theorem,

$$\lim_{n\to\infty} \mathrm{P}\left\{X_n \le a\right\} = \mathrm{P}\left\{X \le a\right\}, \tag{20}$$

at all a where F_X is continuous. But X is discrete and integer-valued. Therefore, F_X is continuous at a if and only if a is not a nonnegative integer. If a *is* a nonnegative integer, then we can choose a non-integer $b \in (a\,, a+1)$ to find that

$$\lim_{n\to\infty} \mathrm{P}\{X_n \le b\} = \mathrm{P}\{X \le b\}.$$

Because X_n and X are both non-negative integers, $X_n \le b$ if and only if $X_n \le a$, and $X \le b$ if and only if $X \le a$. Therefore, (20) holds for all a.

Example 27.4 (The de Moivre–Laplace central limit theorem)**.** Suppose $X_n = \text{binomial}(n\,,p)$, where $p \in (0\,,1)$ is fixed, and define Y_n to be its *standardization*. That is, $Y_n = (X_n - \mathrm{E}X_n)/\sqrt{\mathrm{Var}X_n}$. Alternatively,

$$Y_n = \frac{X_n - np}{\sqrt{npq}}.$$

We know that for all real numbers t,

$$M_{X_n}(t) = \left(q + pe^{-t}\right)^n.$$

We can use this to compute M_{Y_n} as follows:

$$M_{Y_n}(t) = \mathrm{E}\left[\exp\left(t \cdot \frac{X_n - np}{\sqrt{npq}}\right)\right].$$

Recall that $X_n = I_1 + \cdots + I_n$, where I_j is one if the jth trial succeeds; else, $I_j = 0$. Then, $I_1, \ldots, I_n$ are independent binomial$(1\,,p)$'s, and $X_n - np = \sum_{j=1}^n (I_j - p)$. Therefore,

$$\begin{aligned}
\mathrm{E}\left[\exp\left(t \cdot \frac{X_n - np}{\sqrt{npq}}\right)\right] &= \mathrm{E}\left[\frac{t}{\sqrt{npq}} \sum_{j=1}^n (I_j - p)\right] \\
&= \left(\mathrm{E}\left[\exp\left(\frac{t}{\sqrt{npq}}(I_1 - p)\right)\right]\right)^n \\
&= \left(p \exp\left\{\frac{t}{\sqrt{npq}}(1-p)\right\} + q\exp\left\{\frac{t}{\sqrt{npq}}(0-p)\right\}\right)^n \\
&= \left(p\exp\left\{t\sqrt{\frac{q}{np}}\right\} + q\exp\left\{-t\sqrt{\frac{p}{nq}}\right\}\right)^n.
\end{aligned}$$

According to the Taylor–MacLaurin expansion,

$$\exp\left\{t\sqrt{\frac{q}{np}}\right\} = 1 + t\sqrt{\frac{q}{np}} + \frac{t^2 q}{2np} + \text{smaller terms},$$
$$\exp\left\{-t\sqrt{\frac{p}{nq}}\right\} = 1 - t\sqrt{\frac{p}{nq}} + \frac{t^2 p}{2nq} + \text{smaller terms}.$$

Therefore,

$$\begin{aligned}
&p\exp\left\{t\sqrt{\frac{q}{np}}\right\} + q\exp\left\{-t\sqrt{\frac{p}{nq}}\right\}\\
&= p\left(1 + t\sqrt{\frac{q}{np}} + \frac{t^2 q}{2np} + \cdots\right) + q\left(1 - t\sqrt{\frac{p}{nq}} + \frac{t^2 p}{2nq} + \cdots\right)\\
&= p + t\sqrt{\frac{pq}{n}} + \frac{t^2 q}{2n} + \cdots + q - t\sqrt{\frac{pq}{n}} + \frac{t^2 p}{2n} + \cdots\\
&= 1 + \frac{t^2}{2n} + \text{smaller terms}.
\end{aligned}$$

Consequently,

$$M_{Y_n}(t) = \left(1 + \frac{t^2}{2n} + \text{smaller terms}\right)^n \to \exp\left(-\frac{t^2}{2}\right).$$

We recognize the right-hand side as $M_Y(t)$, where $Y = N(0\,,1)$. Because F_Y is continuous, this prove the *central limit theorem* of de Moivre: For all real numbers a,

$$\lim_{n\to\infty} \mathrm{P}\{Y_n \le a\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-x^2/2}\,dx.$$

2. Jointly distributed continuous random variables

Definition 27.5. We say that (X, Y) is jointly distributed with *joint density function* f if f is piecewise continuous, and for all "nice" two-dimensional sets A,

$$\mathrm{P}\{(X,Y) \in A\} = \iint_A f(x\,,y)\,dx\,dy.$$

If (X, Y) has a joint density function f, then:

(1) $f(x\,,y) \ge 0$ for all x and y;

(2) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x\,,y)\,dx\,dy = 1$.

Any function f of two variables that satisfies these properties will do.

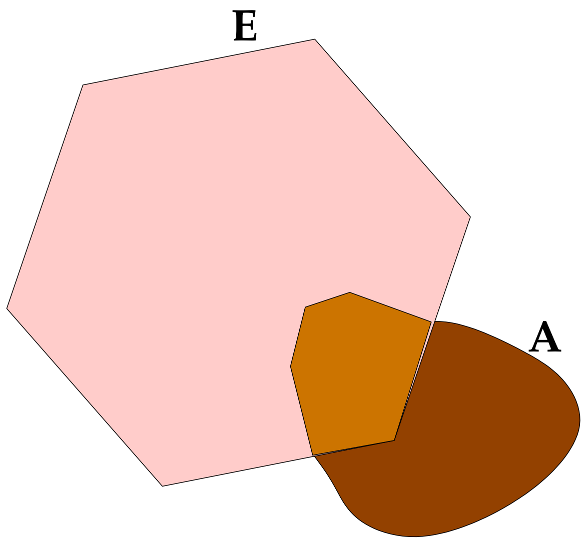

Figure 1. Region of integration in Example 27.6

Example 27.6 (Uniform joint density)**.** Suppose A is a subset of the plane that has a well-defined finite area $|A| > 0$. Define

$$f(x\,,y) = \begin{cases} \dfrac{1}{|A|} & \text{if } (x\,,y) \in A, \\ 0 & \text{otherwise.} \end{cases}$$

Then, f is a joint density function, and the corresponding random vector (X, Y) is said to be distributed *uniformly* on A. Moreover, for all planar sets E with well-defined areas,

$$\mathrm{P}\{(X, Y) \in E\} = \iint_{E \cap A} \frac{1}{|A|}\, dx\, dy = \frac{|E \cap A|}{|A|}.$$

See Figure 1.

Example 27.7. Suppose (X, Y) has joint density

$$f(x\,,y) = \begin{cases} Cxy & \text{if } 0 < y < x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

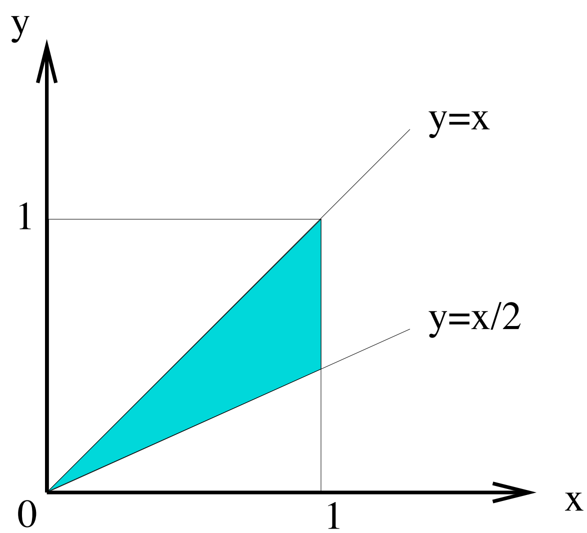

Figure 2. Region of integration in Example 27.7

Let us first find C, and then $\mathrm{P}\{X \le 2Y\}$. To find C:

$$1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x\,, y)\, dx\, dy = \int_0^1 \int_0^x Cxy\, dy\, dx$$
$$= C \int_0^1 x \underbrace{\left(\int_0^x y\, dy \right)}_{\frac{1}{2}x^2} dx = \frac{C}{2} \int_0^1 x^3\, dx = \frac{C}{8}.$$

Therefore, $C = 8$, and hence

$$f(x\,, y) = \begin{cases} 8xy & \text{if } 0 < y < x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Now

$$\mathrm{P}\{X \le 2Y\} = \mathrm{P}\{(X, Y) \in A\} = \iint_A f(x\,, y)\, dx\, dy,$$

where A denotes the collection of all points $(a\,, b)$ in the plane such that $a \le 2b$. Therefore,

$$\mathrm{P}\{X \le 2Y\} = \int_0^1 \int_{x/2}^x 8xy\, dy\, dx = \frac{3}{32}.$$

See Figure 2.