ECE 361 Probability for Engineers (Fall, 2016)
Lecture 5a

---

§2.7 Independence

Independence of several RVs

We say RVs X, Y, Z are independent if their joint PMF factors as the product of the marginal PMFs

$$p_{X,Y,Z}(x, y, z) = p_X(x)p_Y(y)p_Z(z), \ \forall (x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}. \tag{1}$$

Variance of the sum of independent RVs

If $X_1, \ldots, X_n$ are independent RVs and $a_1, \ldots, a_n$ are scalars then

$$\boxed{\mathrm{var}(a_1X_1 + \cdots + a_nX_n) = a_1^2\mathrm{var}(X_1) + \cdots + a_n^2\mathrm{var}(X_n).} \tag{2}$$

Example. Variance of the binomial and the Poisson. Let $X_1, \ldots, X_n$ be independent Bernoulli RVs with success probability p and let $X = X_1 + \cdots + X_n$ be the number of successes, so $X \sim \mathrm{Ber}(n, p)$. Then

$$\mathrm{var}(X) = \mathrm{var}(X_1 + \cdots + X_n) = n\mathrm{var}(X_1) = np(1 - p). \tag{3}$$

Recall the Poisson RV $Y \sim \mathrm{Po}(\lambda)$ was seen as a limit of the binomial RV $X \sim \mathrm{Bin}(n, p)$ when $n \to \infty$ and $p = p(n) \to 0$ such that $\lim_{n\to\infty} np(n) = \lambda$. In this way we see that $\mathrm{var}(Y) = \lim_{n\to\infty} np(n)(1 - p(n)) = \lambda$. Recall $\mathbb{E}[Y] = \lambda$ as well.

Summary of important discrete distributions

- Bernoulli: model a single random trial that either succeeds (with probability p) or fails (with probability $1 - p$)
  - Notation: $X \sim \mathrm{Ber}(p)$, $p \in [0, 1]$
  - Support: $\mathcal{X} = \{0, 1\}$
  - Distribution: $p_X(0) = 1 - p$, $p_X(1) = p$
  - Mean: $\mathbb{E}[X] = p$
  - Variance: $\mathrm{Var}(X) = p(1 - p)$
- binomial: model of the number of successes in n independent Bernoulli trials
  - Notation: $X \sim \mathrm{Bin}(n, p)$, $p \in [0, 1]$, $n \in \mathbb{N}$
  - Support: $\mathcal{X} = \{0, \ldots, n\}$
  - Distribution: $p_X(k) = \binom{n}{k} p^k (1 - p)^{n-k}$
  - Mean: $\mathbb{E}[X] = np$
  - Variance: $\mathrm{Var}(X) = np(1 - p)$
- geometric: model of the number of independent Bernoulli trials required until the first success
  - Notation: $X \sim \mathrm{Geo}(p)$, $p \in [0, 1]$
  - Support: $\mathcal{X} = \{1, 2, 3, \ldots\}$
  - Distribution: $p_X(k) = (1 - p)^{k-1}p$
  - Mean: $\mathbb{E}[X] = 1/p$
  - Variance: $\mathrm{Var}(X) = (1 - p)/p^2$
- Poisson: limiting distribution for the binomial distribution when $p(n)$ scales such that $np(n) \to \lambda$ as $n \to \infty$
  - Notation: $X \sim \mathrm{Po}(\lambda)$, $\lambda \geq 0$
  - Support: $\mathcal{X} = \{0, 1, 2, 3, \ldots\}$

  - Distribution: $p_X(k) = \mathrm{e}^{-\lambda}\lambda^k/k!$
  - Mean: $\mathbb{E}[X] = \lambda$
  - Variance: $\mathrm{Var}(X) = \lambda$

- uniform: (discrete) uniform distribution over outcomes 1 through n
  - Notation: $X \sim \mathrm{Uni}[1, n]$
  - Support: $\mathcal{X} = \{1, \ldots, n\}$
  - Distribution: $p_X(k) = 1/n$
  - Mean: $\mathbb{E}[X] = (n+1)/2$
  - Variance: $\mathrm{Var}(X) = (n^2 - 1)/12$

§3.1 Continuous RVs and their PDFs

We now begin the material in Chapter 3: general RVs. In Chapter 2 we covered Discrete RVs; recall X is a discrete RV if its support $\mathcal{X}$ is finite or countably infinite. Now we introduce Continuous RVs and then later on Mixed RVs.

An RV X is continuous if there exists a non-negative function $f_X : \mathcal{X} \to \mathbb{R}_+$ called the probability density function (PDF) such that:

$$\mathbb{P}(X \in B) = \int_B f_X(x)\mathrm{d}x, \ \forall B \subseteq \mathbb{R}. \tag{4}$$

In particular for the case of intervals, $B = [a, b]$:

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x)\mathrm{d}x. \tag{5}$$

See Fig. 3.1 for a helpful illustration. Because the support of a continuous RV is uncountably infinite it follows that the probability of any given value is zero:

$$\mathbb{P}(X = a) = \int_a^a f_X(x)\mathrm{d}x = 0. \tag{6}$$

Thus the inclusion or exclusion of endpoints of intervals has no effect, i.e., $[a, b), [a, b], (a, b], (a, b)$ have the same probability. Besides non-negativity, a PDF must also be normalized:

$$\mathbb{P}(X \in \mathbb{R}) = \int_{-\infty}^{\infty} f_X(x)\mathrm{d}x = 1. \tag{7}$$

We interpret the PDF as a "probability mass per unit length":

$$\mathbb{P}(X \in [x, x+\delta]) = \int_x^{x+\delta} f_X(x)\mathrm{d}x \approx f_X(x)\delta \ \Rightarrow \ f_X(x) \approx \mathbb{P}(X \in [x, x+\delta])/\delta. \tag{8}$$

See Fig. 3.2 for a helpful illustration.

Example. Continuous uniform RV. A gambler spins a wheel continuously calibrated between 0 and 1 such that any two sub-intervals of the same length have the same probability. Thus

$$f_X(x) = \begin{cases} c, & 0 \leq x \leq 1 \\ 0, & \text{else} \end{cases} \tag{9}$$

for some $c \in \mathbb{R}_+$. The normalization property requires:

$$\int_0^1 f_X(x)\mathrm{d}x = c = 1. \tag{10}$$

If instead $\mathcal{X} = [a, b]$ then we find $f_X(x) = 1/(b-a)$ for $x \in [a, b]$.

Example. Piecewise constant PDF. Alvin's driving time to work is uniformly random between 15 and 20 when sunny and uniformly random between 20 and 25 when rainy. Suppose a day is sunny with probability 2/3 and rainy with probability 1/3. What is the PDF of the driving time RV X? Let $\mathcal{X} = [15, 25]$ be the support for X and note by construction:

$$f_X(x) = \begin{cases} c_1, & 15 \leq x \leq 20 \\ c_2, & 20 \leq x \leq 25 \end{cases}, \tag{11}$$

and we find c_1, c_2 as follows:

$$\begin{aligned} \frac{2}{3} &= \mathbb{P}(\text{sunny}) = \int_{15}^{20} f_X(x)\mathrm{d}x = 5c_1 \\ \frac{1}{3} &= \mathbb{P}(\text{rainy}) = \int_{20}^{25} f_X(x)\mathrm{d}x = 5c_2 \end{aligned} \tag{12}$$

which yields $c_1 = 2/15$ and $c_2 = 1/15$.

Example. A PDF that can take arbitrarily large values. Consider

$$f_X(x) = \begin{cases} \frac{1}{2\sqrt{x}}, & 0 < x \leq 1 \\ 0, & \text{else} \end{cases}. \tag{13}$$

This function grows arbitrarily large as $x \to 0$, but still is non-negative and integrates to one:

$$\int_{-\infty}^{\infty} f_X(x)\mathrm{d}x = \int_0^1 \frac{1}{2\sqrt{x}}\mathrm{d}x = \sqrt{x}\Big|_0^1 = 1. \tag{14}$$

Summary of PDF properties. Let X be a continuous RV with PDF f_X.

- $f_X(x) \geq 0$ for all x.
- $\int_{-\infty}^{\infty} f_X(x)\mathrm{d}x = 1$.
- If δ is very small then $\mathbb{P}(X \in [x, x+\delta]) \approx f_X(x)\delta$.
- For any $B \subseteq \mathbb{R}$: $\mathbb{P}(X \in B) = \int_B f_X(x)\mathrm{d}x$.

References

[1] *Introduction to Probability, 2nd Edition* by Dimitri P. Bertsekas and John N. Tsitsiklis, Athina Scientific Press, 2008.