Recitation Notes for PSTAT 120B

Haosheng Zhou

Jan, 2023

Week 1

This week we will review some of the important properties taught in 120A as a preparation for homework 0.

The first concept to review is the **continuous and discrete random variables**. Generally, a random variable

$$X(\omega) : \Omega \to \mathbb{R} \tag{1}$$

is a mapping from the sample space to the real numbers, i.e. it assigns a value to each possible outcome of random experiment. Discrete random variables can take countably many values while continuous random variables can take uncountably many values. For example, if we want to consider the random variable X as the outcome after rolling one dice, then we have to first specify the **sample space**, i.e. the set of all possible outcomes rolling one dice, which should be $\Omega = \{1, 2, ..., 6\}$. As a result, such random variable X is defined as

$$X(\omega) = \omega \tag{2}$$

an identity map. Since X can only take values in $\{1, 2, ..., 6\}$, a finite set, it's a discrete random variable.

To describe a single random variable, we have the **cumulative distribution function (CDF)** defined for any random variable X as

$$F(x) = \mathbb{P}\left(X \leq x\right) \tag{3}$$

Such F is always right-continuous, increasing and $F(-\infty) = 0, F(+\infty) = 1$ (try to explain the meaning of those properties). In particular, for continuous random variable such F is continuous and for discrete random variable such F is a step function. For continuous random variables, assume that F is nice enough to be differentiable so $F' = f$ gives the **density** that characterizes the distribution of the continuous random variable (for random vectors, those concepts can be generalized).

To describe the relationship between two random variables, the most important property is **independence**. We call X, Y independent if

$$\forall x, y \in \mathbb{R}, \mathbb{P}\left(X \leq x, Y \leq y\right) = \mathbb{P}\left(X \leq x\right)\mathbb{P}\left(Y \leq y\right) \tag{4}$$

which can also be explained in the sense of conditional probability (try to write the equality in the conditional form). For discrete r.v. X, Y, they are independent if and only if $\forall x, y \in \mathbb{R}, \mathbb{P}\left(X = x, Y = y\right) = \mathbb{P}\left(X = x\right)\mathbb{P}\left(Y = y\right)$ and for continuous r.v. X, Y, they are independent if and only if $f_X(x)f_Y(y) = f_{X,Y}(x, y)$ *a.e.* (think about why the criterion for discrete r.v. does not hold for continuous r.v.).

The important concept to mention is the **expectation** of continuous or discrete random variables. For discrete random variable X, assume that its distribution is given by

$$p_k = \mathbb{P}\left(X = a_k\right) \ \ (k = 0, 1, ...) \tag{5}$$

so the expectation is formed as

$$\mathbb{E}X = \sum_{k=0}^{\infty} a_k \cdot \mathbb{P}\left(X = a_k\right) = \sum_{k=0}^{\infty} a_k \cdot p_k \tag{6}$$

i.e., the **sum** of the product of the possible value a_k taken by X and the probability of X taking value a_k.

For discrete random variable X, assume that its density is $f(x)$, so the expectation is formed as

$$\mathbb{E}X = \int_{\mathbb{R}} x f(x)\, dx \tag{7}$$

i.e., the **integral** of the product of the possible value x taken by X and $f(x)$, the likelihood of X taking value x. In the homework, we will be asked to prove the **linearity of expectation** by using those definitions.

Another important concept is the **variance**, defined as

$$Var(X) = \mathbb{E}(X - \mathbb{E}X)^2 \tag{8}$$

the connection between variance and expectation can be given by the useful formula that

$$Var(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2 \tag{9}$$

for two random variables, we can define the **covariance** to describe their relationship

$$cov(X, Y) = \mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)] \tag{10}$$

and a similar identity holds that

$$cov(X) = \mathbb{E}XY - (\mathbb{E}X)(\mathbb{E}Y) \tag{11}$$

note that $cov(X, X) = Var(X)$ and that $cov(X, Y)$ is **bilinear**, i.e. $cov(aX + bY, Z) = a \cdot cov(X, Z) + b \cdot cov(Y, Z), cov(Z, aX + bY) = a \cdot cov(Z, X) + b \cdot cov(Z, Y)$ and **symmetric**, i.e. $cov(X, Y) = cov(Y, X)$. This is especially useful when computing the variance of a linear combination. For example, if we want to write $Var(2X + 3Y)$ in terms of $Var(X), Var(Y)$,

$$\begin{aligned}
Var(2X + 3Y) &= cov(2X + 3Y, 2X + 3Y) && (12)\\
&= 2cov(X, 2X + 3Y) + 3cov(Y, 2X + 3Y) && (13)\\
&= 2[2cov(X, X) + 3cov(X, Y)] + 3[2cov(Y, X) + 3cov(Y, Y)] && (14)\\
&= 4Var(X) + 12cov(X, Y) + 9cov(Y, Y) && (15)
\end{aligned}$$

you are asked to prove a more general version of this property in the homework.

Finally, let's talk about **normal distribution**. We say $X \sim N(\mu, \sigma^2)$ if it has density

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \ (x \in \mathbb{R}) \tag{16}$$

for the two parameters μ, σ^2 of normal distribution, a direct interpretation is that $\mathbb{E}X = \mu, Var(X) = \sigma^2$. You can try to prove those properties on your own by applying the definitions of expectation and variance to calculate the integrals. A trick will be that when calculating the integral

$$\mathbb{E}X = \int_{\mathbb{R}} x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx \tag{17}$$

use the change of variables $u = \frac{x-\mu}{\sigma}$ to make the life easier

$$\begin{aligned}
\mathbb{E}X &= \int_{\mathbb{R}} x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx && (18)\\
&= \frac{1}{\sqrt{2\pi}\sigma} \sigma \int_{\mathbb{R}} (\sigma u + \mu) e^{-\frac{u^2}{2}} \, du && (19)\\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (\sigma u + \mu) e^{-\frac{u^2}{2}} \, du && (20)\\
&= \frac{\mu}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{u^2}{2}} \, du && (21)\\
&= \mu && (22)
\end{aligned}$$

here we use the property that $ue^{-\frac{u^2}{2}}$ is an odd function and that $\int_{\mathbb{R}} e^{-\frac{u^2}{2}} \, du = \sqrt{2\pi}$ (this property can be deduced from the standard normal density, en easy way to remember). The calculation of variance is left to the reader.

The **standard normal CDF** is one of the most frequently used notations in statistics. The definition is

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \, dt = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \, dt = \mathbb{P}(G \leq x) \ \ (G \sim N(0, 1)) \tag{23}$$

a property of Φ is that

$$\forall x \in \mathbb{R}, \Phi(x) + \Phi(-x) = 1 \tag{24}$$

to see this, notice that $\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$ is an even function in t, so

$$\begin{aligned}
\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \, dt &= \int_{-x}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \, du \ (u = -t) && (25)\\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \, du - \int_{-\infty}^{-x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \, du && (26)\\
&= 1 - \Phi(-x) && (27)
\end{aligned}$$