556: MATHEMATICAL STATISTICS I
EXPECTATIONS

- Random variable X
- Mass/density function f_X with support $\mathbb{X}$.
- **Expectation**

$$E_{f_X}[X] = \begin{cases} \displaystyle\sum_{x \in \mathbb{X}} x f_X(x) & X \text{ discrete} \\ \displaystyle\int_{-\infty}^{\infty} x f_X(x) dx \quad = \int_{\mathbb{X}} x f_X(x) dx & X \text{ continuous} \end{cases}$$

In the discrete case, if X only takes values on (a subset of) the integers, we can also write

$$E_{f_X}[X] = \sum_{x=-\infty}^{\infty} x f_X(x)$$

- **Extension**: Let g be a real-valued function whose domain includes X. Then

$$E_{f_X}[g(X)] = \begin{cases} \displaystyle\sum_{x=-\infty}^{\infty} g(x) f_X(x) & X \text{ discrete} \\ \displaystyle\int_{-\infty}^{\infty} g(x) f_X(x) dx & X \text{ continuous} \end{cases}$$

Note that the sum/integral may be **divergent**, so that the expectation is **not finite**.

All definitions, and the following properties, extend to the **vector** random variable case.

PROPERTIES

1 **Linearity:** Let g and h be real-valued functions whose domains include $\mathbb{X}$, and let a and b be constants.

$$\begin{aligned} E_{f_X}[ag(X) + bh(X)] &= \int [ag(x) + bh(x)] f_X(x) dx \\ &= a \int g(x) f_X(x) dx + b \int h(x) f_X(x) dx \\ &= a E_{f_X}[g(X)] + b E_{f_X}[h(X)] \end{aligned}$$

Hence, for example,

$$E_{f_X}[aX + b] = a E_{f_X}[X] + b.$$

2 Let $\mu = E_{f_X}[X]$, and consider $g(x) = (x - \mu)^2$. Then

$$\begin{aligned} E_{f_X}[g(X)] &= \int (x-\mu)^2 f_X(x) dx = \int x^2 f_X(x) dx - 2\mu \int x f_X(x) dx + \mu^2 \int f_X(x) dx \\ &= \int x^2 f_X(x) dx - 2\mu^2 + \mu^2 = \int x^2 f_X(x) dx - \mu^2 \\ &= E_{f_X}[X^2] - \{E_{f_X}[X]\}^2 \end{aligned}$$

Thus

(i) **Variance:** $Var_{f_X}[X] = E_{f_X}[X^2] - \{E_{f_X}[X]\}^2$

(ii) **Standard deviation** $\sqrt{Var_{f_X}[X]}$

3 Consider $g(x) = x^r$ for $r = 1, 2, ...$ Then in the continuous case

$$E_{f_X}[g(X)] = E_{f_X}[X^r] = \int x^r f_X(x)dx,$$

and $E_{f_X}[X^r]$ is the rth **moment** of the distribution.

4 Consider $g(x) = (x - \mu)^r$ for $r = 1, 2, ...$ Then

$$E_{f_X}[g(X)] = E_{f_X}[(X - \mu)^r] = \int (x - \mu)^r f_X(x)dx,$$

and $E_{f_X}[(X - \mu)^r]$ is the rth **central moment** of the distribution.

5 Consider $g(x) = aX + b$. Then

$$\begin{aligned} Var_{f_X}[g(X)] = E_{f_X}[(aX + b - E_{f_X}[aX + b])^2] &= E_{f_X}[(aX + b - aE_{f_X}[X] - b)^2] \\ &= E_{f_X}[(a^2(X - E_{f_X}[X])^2] \\ &= a^2 Var_{f_X}[X]. \end{aligned}$$

so

$$Var_{f_X}[aX + b] = a^2 Var_{f_X}[X].$$

6 Consider $g(x) = e^{tx}$, for constant $t \in (-h, h)$ for some $h > 0$, and

$$M_X(t) = E_{f_X}[g(X)] = E_{f_X}\left[e^{tX}\right].$$

Then $M_X(t)$ is the **moment generating function**.

7 Consider $K_X(t) = \log M_X(t)$. Then $K_X(t)$ is the **cumulant generating function**.

8 Consider $g(x) = e^{itx}$, where $i = \sqrt{-1}$.

$$C_X(t) = E_{f_X}[g(X)] = E_{f_X}\left[e^{itX}\right].$$

Then $C_X(t)$ is the **characteristic function**.

In the discrete case, for each of these properties, we replace integrals by sums.

Note that in the vector random variable case, generating functions have vector arguments. For example, the joint mgf for vector r.v. $\underset{\sim}{X} = (X_1, \ldots, X_k)^\mathsf{T}$ is a function of $\underset{\sim}{t} = (t_1, \ldots, t_k)^\mathsf{T}$

$$M_{\underset{\sim}{X}}(\underset{\sim}{t}) = E_{f_{\underset{\sim}{X}}}\left[\exp\left\{\underset{\sim}{t}^\mathsf{T}\underset{\sim}{X}\right\}\right] = E_{f_{\underset{\sim}{X}}}\left[\exp\left\{\sum_{j=1}^{k} t_j X_j\right\}\right]$$