# STA 103 (SS2 2025): Formula Sheet 2
## Instructor: Wookyeong Song
## Institution: UC Davis

- Expectation formula:

$$E(aX+b) = aE(X) + b$$
$$E(aX+bY) = aE(X) + bE(Y)$$

- Variance formula:

$$\mathrm{var}(aX+b) = \mathrm{var}(aX) = a^2\,\mathrm{var}(X)$$
$$\mathrm{var}(aX+bY) = a^2\mathrm{var}(X) + b^2\mathrm{var}(Y) + 2ab\,\mathrm{cov}(X,Y)$$
$$\mathrm{var}(aX-bY) = a^2\mathrm{var}(X) + b^2\mathrm{var}(Y) - 2ab\,\mathrm{cov}(X,Y)$$

- Covariance properties

$$\mathrm{cov}(aX,bY) = a\cdot b\cdot \mathrm{cov}(X,Y)$$
$$\mathrm{cov}(X,c) = \mathrm{cov}(c,X) = 0$$
$$X \text{ \& } Y \text{ are independent} \Longrightarrow \mathrm{cov}(X,Y) = 0$$

- Continuous probability density functions (PDF): If $X$ has PDF $p(x)$ then

$$Pr(a \le X \le b) = \int_a^b p(x)\,dx$$
$$E(X) = \int_a^b x\,p(x)\,dx$$
$$E(f(X)) = \int_a^b f(x)\,p(x)\,dx$$

- Uniform Random Variable
  - The PDF for $Y \sim \mathrm{Unif}[a,b]$ is:

$$f(y) = \begin{cases} \frac{1}{(b-a)}, & a \le y \le b, \\ 0, & \text{otherwise.} \end{cases}$$

  - Mean and Variance:

$$E(Y) = \frac{a+b}{2}, \quad \mathrm{Var}(Y) = \frac{(b-a)^2}{12}.$$

- Normal Random Variables
  - The PDF for $X \sim N(\mu,\sigma^2)$ is:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

  - Mean and Variance:

$$E(X) = \mu, \quad \mathrm{Var}(X) = \sigma^2.$$

- $z$-scores: For any random variable $X$ we can change units to $W$ via

$$W = \frac{X-a}{b}$$
$$X = a + bW.$$

If we use $a = E(X)$ and $b = sd(X)$ then $W$ is called a $z$-score transformation of $X$ and satisfies

$$E(W) = 0 \text{ and } \mathrm{sd}(W) = 1.$$

- Special Normal/Gaussian Cases: If $X \sim \mathcal{N}(a,b)$ then

$$E(X) = a$$
$$sd(X) = \sqrt{b}$$
$$W = \frac{X-a}{\sqrt{b}} \sim N(0,1)$$

- The central limit theorem (CLT): If $X_1,\ldots,X_n$ are independent random variables, all with the same PMF or PDF, then

$$\bar{X} = \frac{X_1+\cdots+X_n}{n} \approx N(a,b)$$

for large $n$ (usually $n > 30$ is good), where $a = E(\overline{X})$ and $b = \mathrm{Var}(\overline{X})$.

- For i.i.d. RVs $X_1,\ldots,X_n$, with $E(X_1) = \mu$ and $\mathrm{Var}(X_1) = \sigma^2$.

$$E(\bar{X}) = E(X_1) = \mu, \quad \mathrm{Var}(\bar{X}) = \frac{\mathrm{Var}(X_1)}{n} = \frac{\sigma^2}{n}.$$

- Bernoulli Distribution
  1. $X \sim \mathrm{Bernoulli}(p)$, Parameters: $p$.
  2. PMF:

$$p(x) = \begin{cases} p^x(1-p)^{1-x} & \text{if } x = 0 \text{ or } x = 1, \\ 0 & \text{otherwise.} \end{cases}$$

  3. Mean $E(X) = p$, Variance $\mathrm{Var}(X) = p(1-p)$.

- Binomial Distribution $X \sim \mathrm{Binomial}(n,p)$
  1. Parameters: $n$, $p$.
  2. PMF:

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

  3. $\binom{n}{x} = \frac{n!}{(n-x)!x!}$
  4. Mean $E(X) = np$,
     Variance $\mathrm{Var}(X) = np(1-p)$.
  5. Using the CLT, if $np \ge 10$ and $n(1-p) \ge 10$, we can approximate $X \sim N\left(np, np(1-p)\right)$.