CHAPTER 5

Jointly Distributed Random Variables

Joint Probability Mass Function

Let X and Y be two discrete rv's defined on the sample space of an experiment. The *joint probability mass function* $p(x, y)$ is defined for each pair of numbers (x, y) by

$$p(x, y) = P(X = x \text{ and } Y = y)$$

Let A be the set consisting of pairs of (x, y) values, then

$$P\left[(X, Y) \in A\right] = \sum_{(x,y) \in A} \sum p(x, y)$$

Marginal Probability Mass Functions

The *marginal probability mass functions* of X and Y, denoted $p_X(x)$ and $p_Y(y)$ are given by

$$p_X(x) = \sum_y p(x, y) \qquad p_Y(y) = \sum_x p(x, y)$$

- – the height and weight of a person;
- – the temperature and rainfall of a day;
- – the two coordinates of a needle randomly dropped on a table;
- – the number of 1s and the number of 6s in 10 rolls of a die.

- **Example.** We are interested in the effect of seat belt use on saving lives. If we consider the following random variables X_1 and X_2 defined as follows:
- $X_1 = 0$ if child survived
- $X_1 = 1$ if child did not survive

- And $X_2 = 0$ if no belt
- $X_2 = 1$ if adult belt used
- $X_2 = 2$ if child seat used

- The following table represents the ***joint probability distribution*** of X_1 and X_2 . In general we write $P(X_1 = x_1 , X_2 = x_2) = p(x_1 , x_2)$ and call $p(x_1 , x_2)$ the ***joint probability function*** of (X_1 , X_2).

		X_1 = 0	X_1 = 1	
	0	0.38	0.17	0.55
X_2	1	0.14	0.02	0.16
	2	0.24	0.05	0.29
		0.76	0.24	

- Probability that a child will both survive and be in a child seta when involved in an accident is:
- $P(X_1 = 0, X_2 = 2) = 0.24$
- Probability that a child will be in a child seat:
- $P(X_2 = 2) = P(X_1 = 0,\ X_2 = 2) + P(X_1 = 1,\ X_2 = 2) =$ $0.24+0.05= 0.29$

Example: 3 sox in a box (numbered 1,2,3). Draw 2 sox at random w/o replacement. $X =$ # of first sock, $Y =$ # of second sock. The joint pmf $f(x, y)$ is

	$X = 1$	$X = 2$	$X = 3$	$\Pr(Y = y)$
$Y = 1$	0	1/6	1/6	1/3
$Y = 2$	1/6	0	1/6	1/3
$Y = 3$	1/6	1/6	0	1/3
$\Pr(X = x)$	1/3	1/3	1/3	1

$\Pr(X = x)$ is the "marginal" distribution of X.

$\Pr(Y = y)$ is the "marginal" distribution of Y.

By the law of total probability,

$$\Pr(X=1) \;=\; \sum_{y=1}^{3} \Pr(X=1, Y=y) \;=\; 1/3.$$

In addition,

$$\begin{aligned}
&\Pr(X \geq 2, Y \geq 2) \\
&\quad = \sum_{x \geq 2} \sum_{y \geq 2} f(x,y) \\
&\quad = f(2,2) + f(2,3) + f(3,2) + f(3,3) \\
&\quad = 0 + 1/6 + 1/6 + 0 \;=\; 1/3.
\end{aligned}$$

Joint Probability Density Function

Let X and Y be continuous rv's. Then $f(x, y)$ is a *joint probability density function* for X and Y if for any two-dimensional set A

$$P\left[\left(X,Y\right)\in A\right]=\iint_A f(x,y)\,dx\,dy$$

If A is the two-dimensional rectangle $\left\{(x,y): a\le x\le b, c\le y\le d\right\},$

$$P\left[\left(X,Y\right)\in A\right]=\int_a^b\int_c^d f(x,y)\,dy\,dx$$

Marginal Probability Density Functions

The *marginal probability density functions* of X and Y, denoted $f_X(x)$ and $f_Y(y)$, are given by

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\,dy \quad \text{for } -\infty < x < \infty$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\,dx \quad \text{for } -\infty < y < \infty$$

Example: A bank operates a drive-in facility and a walk-up window.

Let X = proportion of time the drive-in is used

Y = proportion of time the walk-up is used

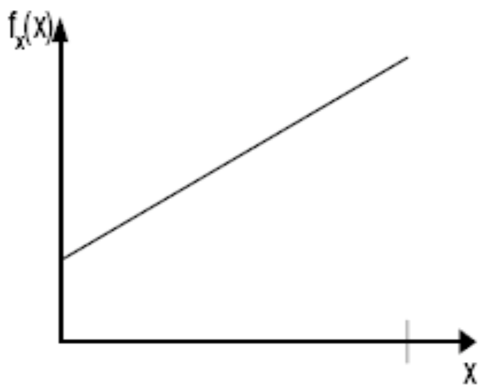
$$f_{XY}(x,y) = \begin{cases} \frac{6}{5}(x+y^2) & \text{for } 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

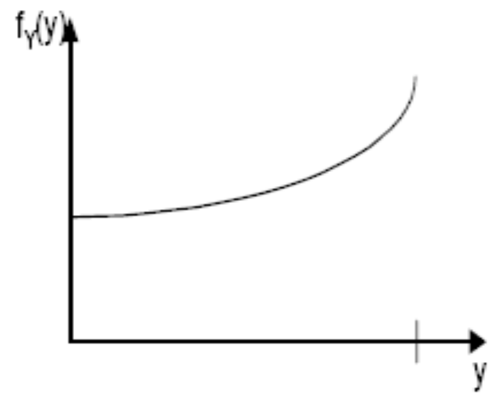
Note: $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dfrac{6}{5}(x+y^2)dxdy = 1$

Marginal density functions:

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x,y)dy = \int_0^1 \frac{6}{5}(x+y^2)dy$$
$$= \begin{cases} \frac{6}{5}(x+\frac{2}{5}) & \text{for } 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x,y)dx = \int_0^1 \frac{6}{5}(x+y^2)dx$$
$$= \begin{cases} \frac{6}{5}(y^2+\frac{3}{5}) & \text{for } 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$





Independent Random Variables

Two random variables X and Y are said to be *independent* if for every pair of x and y values

$$p(x, y) = p_X(x) \cdot p_Y(y)$$

when X and Y are discrete or

$$f(x, y) = f_X(x) \cdot f_Y(y)$$

when X and Y are continuous. If the conditions are not satisfied for all (x, y) then X and Y are *dependent.*

Let the joint density of two random variables x_1 and x_2 be given by

$$f(x_1, x_2) = \begin{cases} 2x_2 e^{-x_1} & x_1 \geq 0, \quad 0 \leq x_2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

First find the marginal density for x_1.

$$\begin{aligned} f_1(x_1) &= \int_0^1 2x_2 \mathrm{e}^{-x_1}\, dx_2 \\ &= x_2^2 \mathrm{e}^{-x_1} \Big|_0^1 \\ &= \mathrm{e}^{-x_1} - 0 \\ &= \mathrm{e}^{-x_1} \end{aligned}$$

Now find the marginal density for x_2.

$$\begin{aligned} f_2(x_2) &= \int_0^\infty 2x_2 \mathrm{e}^{-x_1}\, dx_1 \\ &= -2x_2 \mathrm{e}^{-x_1} \Big|_0^\infty \\ &= 0 - \left(-2x_2 \mathrm{e}^0\right) \\ &= 2x_2 e^0 \\ &= 2x_2 \end{aligned}$$

Conditional Probability Function

Let X and Y be two continuous rv's with joint pdf $f(x, y)$ and marginal X pdf $f_X(x)$. Then for any X value x for which $f_X(x) > 0$, *the conditional probability density function of Y given that $X = x$* is

$$f_{Y|X}(y \mid x) = \frac{f(x, y)}{f_X(x)} \qquad -\infty < y < \infty$$

If X and Y are discrete, replacing pdf's by pmf's gives the *conditional probability mass function of Y when $X = x$*.

Example. Let the joint density of two random variables x and y be given by

$$f(x,y)=\begin{cases}\frac{1}{6}(x+4y) & 0<x<2,\ 0<y<1\\ 0 & \text{otherwise}\end{cases}$$

The marginal density of x is $f_X(x) = \frac{1}{6}(x+2)$ while the marginal density of y is $f_y(y) = \frac{1}{6}(2+8y)$.

Now find the conditional distribution of x given y. This is given by

$$f_{X|Y}(x \mid y) = \frac{f(x,y)}{f(y)} = \frac{\frac{1}{6}(x+4y)}{\frac{1}{6}(2+8y)}$$

$$= \frac{(x+4y)}{(8y+2)}$$

for $0<x<2$ and $0<y<1$. Now find the probability that $X \leq 1$ given that $y = \frac{1}{2}$.

First determine the density function when $y=\frac{1}{2}$ as follows

$$\frac{f(x,y)}{f(y)} = \frac{(x+4y)}{(8y+2)}$$

$$= \frac{\left(x+4(\frac{1}{2})\right)}{\left(8(\frac{1}{2})+2\right)}$$

$$= \frac{(x+2)}{(4+2)} = \frac{(x+2)}{6}$$

Then

$$
\begin{aligned}
P\left(X \le 1 \mid Y = \tfrac{1}{2}\right) &= \int_0^1 \frac{1}{6}(x+2)\,dx \\
&= \frac{1}{6}\left(\frac{x^2}{2} + 2x\right)\Bigg|_0^1 \\
&= \frac{1}{6}\left(\frac{1}{2} + 2\right) - 0 \\
&= \frac{1}{12} + \frac{2}{6} = \frac{5}{12}
\end{aligned}
$$

Let X and Y denote the proportion of two different chemicals in a sample mixture of chemicals used as an insecticide. Suppose X and Y have joint probability density given by:

$$f(x,y)=\begin{cases}2, & 0\le x\le 1, 0\le y\le 1, 0\le x+y\le 1\\ 0, & \textit{elsewhere}\end{cases}$$

(Note that X + Y must be at most unity since the random variables denote proportions within the same sample).

- 1) Find the marginal density functions for X and Y.
- 2) Are X and Y independent?
- 3) Find P(X > 1/2 | Y =1/4).

$$f_1(x)=\begin{cases}\int_0^{1-x} 2dy = 2(1-x), & 0\le x\le 1\\ 0 & \textit{otherwise}\end{cases}$$

$$f_2(y)=\begin{cases}\int_0^{1-y} 2dx = 2(1-y), & 0\le y\le 1\\ 0 & \textit{otherwise}\end{cases}$$

-

- 2) $f_1(x)\, f_2(y)=2(1-x)*\, 2(1-y) \neq 2 = f(x,y)$, for $0 \leq x \leq 1-y$.
- Therefore X and Y are not independent.

- 3)

$$\boldsymbol{P}\left(\boldsymbol{X} > \frac{1}{2} \mid \boldsymbol{Y} = \frac{1}{4}\right) = \int_{1/2}^{1} \boldsymbol{f}(\boldsymbol{x} \mid \boldsymbol{y} = \frac{1}{4})\boldsymbol{dx} = \int_{1/2}^{1} \frac{\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y} = \frac{1}{4})}{\boldsymbol{f}(\boldsymbol{y} = \frac{1}{4})}\boldsymbol{dx} = \int_{1/2}^{1} \frac{2}{2(1-\frac{1}{4})} = \frac{2}{3}$$

Let X and Y be jointly distributed rv's with pmf p(x, y) or pdf f (x, y) according to whether the variables are discrete or continuous. Then the *expected value* of a function $h(X, Y)$, denoted $E[h(X, Y)]$ or $\mu_{h(X,Y)}$

is

$$= \begin{cases} \sum_{x} \sum_{y} h(x, y) \cdot p(x, y) & \text{discrete} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) \cdot f(x, y) dx dy & \text{continuous} \end{cases}$$

Covariance

The *covariance* between two rv's X and Y is

$$\text{Cov}(X,Y) = E\left[(X-\mu_X)(Y-\mu_Y)\right]$$

$$= \begin{cases} \sum_x \sum_y (x-\mu_X)(y-\mu_Y)p(x,y) & \text{discrete} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_X)(y-\mu_Y)f(x,y)\,dx\,dy & \text{continuous} \end{cases}$$

Short-cut Formula for Covariance

$$\mathrm{Cov}(X,Y) = E(XY) - \mu_X \cdot \mu_Y$$

Theorem: X and Y indep implies $\text{Cov}(X,Y) = 0$.

Proof:

$$\begin{aligned} \text{Cov}(X,Y) &= \mathsf{E}[XY] - \mathsf{E}[X]\mathsf{E}[Y] \\ &= \mathsf{E}[X]\mathsf{E}[Y] - \mathsf{E}[X]\mathsf{E}[Y] \ \ (X,\ Y \text{ indep}) \\ &= 0. \end{aligned}$$

Cov(X, Y) = 0 *does not imply* X and Y are independent!!

Definition: The **correlation** between X and Y is

$$\rho = \text{Corr}(X, Y) \equiv \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}.$$

Remark: Cov has "square" units; corr is unitless.

Corollary: X, Y indep implies $\rho = 0$.

Correlation Proposition

1. If a and c are either both positive or both negative, $\text{Corr}(aX + b, cY + d) = \text{Corr}(X, Y)$
2. For any two rv's X and Y,

   $-1 \leq \text{Corr}(X, Y) \leq 1.$

Correlation Proposition

1. If X and Y are independent, then $\rho = 0$, but $\rho = 0$ does not imply independence.
2. $\rho = 1$ or -1 iff $Y = aX + b$ for some numbers a and b with $a \neq 0$.

Theorem: It can be shown that $-1 \leq \rho \leq 1$.

$\rho \approx 1$ is “high” corr

$\rho \approx 0$ is “low” corr

$\rho \approx -1$ is “high” negative corr.

Example: Height is *highly* correlated with weight.

Temperature on Mars has *low* corr with IBM stock price.

Anti-UGA Example: Suppose X is the avg yards/carry that a UGA fullback gains, and Y is his grade on an astrophysics test. Here's the joint pmf $f(x, y)$.

	$X = 2$	$X = 3$	$X = 4$	$f_Y(y)$
$Y = 40$	.0	.2	.1	.3
$Y = 50$	.15	.1	.05	.3
$Y = 60$	.3	.0	.1	.4
$f_X(x)$	.45	.3	.25	1

$$
\begin{aligned}
\mathsf{E}[X] &= \sum_x x f_X(x) = 2.8 \\
\mathsf{E}[X^2] &= \sum_x x^2 f_X(x) = 8.5 \\
\mathsf{Var}(X) &= \mathsf{E}[X^2] - (\mathsf{E}[X])^2 = 0.66
\end{aligned}
$$

Similarly, $\mathsf{E}[Y] = 51$, $\mathsf{E}[Y^2] = 2670$, and $\mathsf{Var}(Y) = 60$.

$$
\begin{aligned}
\mathsf{E}[XY] &= \sum_x \sum_y xy f(x,y) \\
&= 2(40)(.0) + \cdots + 4(60)(.1) = 140 \\
\mathsf{Cov}(X,Y) &= \mathsf{E}[XY] - \mathsf{E}[X]\mathsf{E}[Y] = -2.8 \\
\rho &= \frac{\mathsf{Cov}(X,Y)}{\sqrt{\mathsf{Var}(X)\mathsf{Var}(Y)}} = -0.415.
\end{aligned}
$$

Cts Example: Suppose $f(x, y) = 10x^2y$, $0 \le y \le x \le 1$.

$$f_X(x) = \int_0^x 10x^2 y \, dy = 5x^4, \;\; 0 \le x \le 1$$

$$\mathsf{E}[X] = \int_0^1 5x^5 \, dx = 5/6$$

$$\mathsf{E}[X^2] = \int_0^1 5x^6 \, dx = 5/7$$

$$\mathsf{Var}(X) = \mathsf{E}[X^2] - (\mathsf{E}[X])^2 = 0.01984$$

Similarly,

$$f_Y(y) = \int_y^1 10x^2y\,dx = \frac{10}{3}y(1-y^3),\ \ 0 \le y \le 1$$

$$\mathsf{E}[Y] = 5/9, \quad \mathsf{Var}(Y) = 0.04850$$

$$\mathsf{E}[XY] = \int_0^1 \int_0^x 10x^3y^2\,dy\,dx = 10/21$$

$$\mathsf{Cov}(X,Y) = \mathsf{E}[XY] - \mathsf{E}[X]\mathsf{E}[Y] = 0.1323$$

$$\rho = \frac{\mathsf{Cov}(X,Y)}{\sqrt{\mathsf{Var}(X)\mathsf{Var}(Y)}} = 0.4265$$

Theorem: $\text{Var}(X+Y) = \text{Var}(X)+\text{Var}(Y)+2\text{Cov}(X,Y)$, *whether or not* X and Y are indep.

Remark: If X, Y are indep, the Cov term goes away.

Proof: By the work we did on a previous proof,

$$\begin{aligned} \text{Var}(X+Y) &= \text{E}[X^2] - (\text{E}[X])^2 + \text{E}[Y^2] - (\text{E}[Y])^2 \\ &\quad +2(\text{E}[XY] - \text{E}[X]\text{E}[Y]) \\ &= \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X,Y). \end{aligned}$$

Theorem: $\text{Cov}(aX, bY) = ab\text{Cov}(X, Y)$.

Proof:

$$\begin{aligned}\text{Cov}(aX, bY) &= \text{E}[aX \cdot bY] - \text{E}[aX]\text{E}[bY] \\ &= ab\text{E}[XY] - ab\text{E}[X]\text{E}[Y] \\ &= ab\text{Cov}(X, Y).\end{aligned}$$

Example: $\mathrm{Var}(X-Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)-2\mathrm{Cov}(X,Y)$.

Example:

$$\begin{aligned}&\mathrm{Var}(X-2Y+3Z)\\&\quad=\ \mathrm{Var}(X)+4\mathrm{Var}(Y)+9\mathrm{Var}(Z)\\&\qquad\qquad-4\mathrm{Cov}(X,Y)+6\mathrm{Cov}(X,Z)-12\mathrm{Cov}(Y,Z).\end{aligned}$$

The proportions X and Y of two chemicals found in samples of an insecticide have the joint probability density function

$$f(x, y) = \begin{cases} 2, & 0 \le x \le 1, 0 \le y \le 1, 0 \le x + y \le 1 \\ 0, & \textit{elsewhere} \end{cases}$$

The random variable Z=X + Y denotes the proportion of the insecticide due to both chemicals combined.

1) Find E(Z) and V(Z)

2) Find an interval in which values of Z should lie at least 50% of the samples of insecticide.

3) Find the correlation between X and Y and interpret its meaning.

$$E(X+Y)=\int_0^1\int_0^{1-x}(x+y)2dy\,dx=\int_0^1(1-x^2)dx=\frac{2}{3}$$

$$E[(X+Y)^2]=\int_0^1\int_0^{1-x}(x+y)^2 2dy\,dx=\int_0^1\frac{2}{3}(1-x)^3dx=\frac{2}{3}(x-\frac{x^4}{4})|_0^1=\frac{2}{3}\left(\frac{3}{4}\right)=\frac{1}{2}$$

$$V(Z)=E(Z^2)-E(Z)^2=\frac{1}{2}-\left(\frac{2}{3}\right)^2=\frac{1}{1}$$

Using Tchebysheff's theorem with k= $\sqrt{2}$ we have,

$$\boldsymbol{P}\left(\frac{2}{3}-\sqrt{\frac{2}{18}}<\boldsymbol{X}+\boldsymbol{Y}<\frac{2}{3}+\sqrt{\frac{2}{18}}\right)\geq 0.5$$

The desired interval is (1/3, 1).

$$\rho(X,Y)=\frac{\mathbf{Cov}(X,Y)}{\sqrt{\mathbf{Var}(X)\mathbf{Var}(Y)}}$$

$$f(X) = \int_0^{1-X} f(X,Y)dy = \int_0^{1-X} 2dY = 2(1-X), 0 \le X \le 1$$

$$f(Y) = \int_0^{1-Y} f(X,Y)dx = \int_0^{1-Y} 2dX = 2(1-Y), 0 \le Y \le 1$$

$$E(X) = E(Y) = \int_0^1 z2(1-z)dz = 2\int_0^1 z - z^2 dz = 2\left(\frac{z^2}{2} - \frac{z^3}{3}\right)\Bigg|_0^1 = 2\left(\frac{1}{2} - \frac{1}{3}\right) = 2\left(\frac{1}{6}\right) = \frac{1}{3}$$

$$E(X^2) = E(Y^2) = \int_0^1 z^2(1-z)dz = 2\int_0^1 z^2 - z^3 dz = 2\left(\frac{z^3}{3} - \frac{z^4}{4}\right)\Bigg|_0^1 = 2\left(\frac{1}{3} - \frac{1}{4}\right) = 2\left(\frac{1}{12}\right) = \frac{1}{6}$$

$$Var(X) = Var(Y) = E(X^2) - (E(X))^2 = \frac{1}{6} - \left(\frac{1}{3}\right)^2 = \frac{1}{6} - \frac{1}{9} = \frac{1}{18}$$

$$E(XY) = \int_0^1 \int_0^{1-x} xy2dxdy = \int_0^1 x \int_0^{1-x} 2ydydx = \int_0^1 x(y^2)\Big|_0^{1-x} dx =$$

$$= \int_0^1 x(1-x)^2 dx = \int_0^1 x - 2x^2 + x^3 dx = \left(\frac{x^2}{2} - \frac{2x^3}{3} + \frac{x^4}{4}\right)\Bigg|_0^1 = \frac{1}{2} - \frac{2}{3} + \frac{1}{4} = \frac{1}{12}$$

$$\rho = \frac{\frac{1}{12} - \left(\frac{1}{3}\right)^2}{\sqrt{\left(\frac{1}{18}\right)^2}} = \frac{\frac{1}{12} - \frac{1}{9}}{\frac{1}{18}} = -\frac{1}{2}$$

A *statistic* is any quantity whose value can be calculated from sample data. Prior to obtaining data, there is uncertainty as to what value of any particular statistic will result. A statistic is a random variable denoted by an uppercase letter; a lowercase letter is used to represent the calculated or observed value of the statistic.

Random Samples

The rv's $X_1, \ldots, X_n$ are said to form a (simple *random sample* of size n if

1. The X_i's are independent rv's.
2. Every X_i has the same probability distribution.

Simulation Experiments

The following characteristics must be specified:

1. The statistic of interest.
2. The population distribution.
3. The sample size n.
4. The number of replications k.

Using the Sample Mean

Let $X_1, \ldots, X_n$ be a random sample from a distribution with mean value μ and standard deviation σ. Then

1. $E(\overline{X}) = \mu_{\overline{X}} = \mu$

2. $V(\overline{X}) = \sigma^2_{\overline{X}} = \sigma^2 / n$

In addition, with $T_o = X_1 + \ldots + X_n$,
$E(T_o) = n\mu,\ V(T_o) = n\sigma^2$, and $\sigma_{T_o} = \sqrt{n}\sigma$.

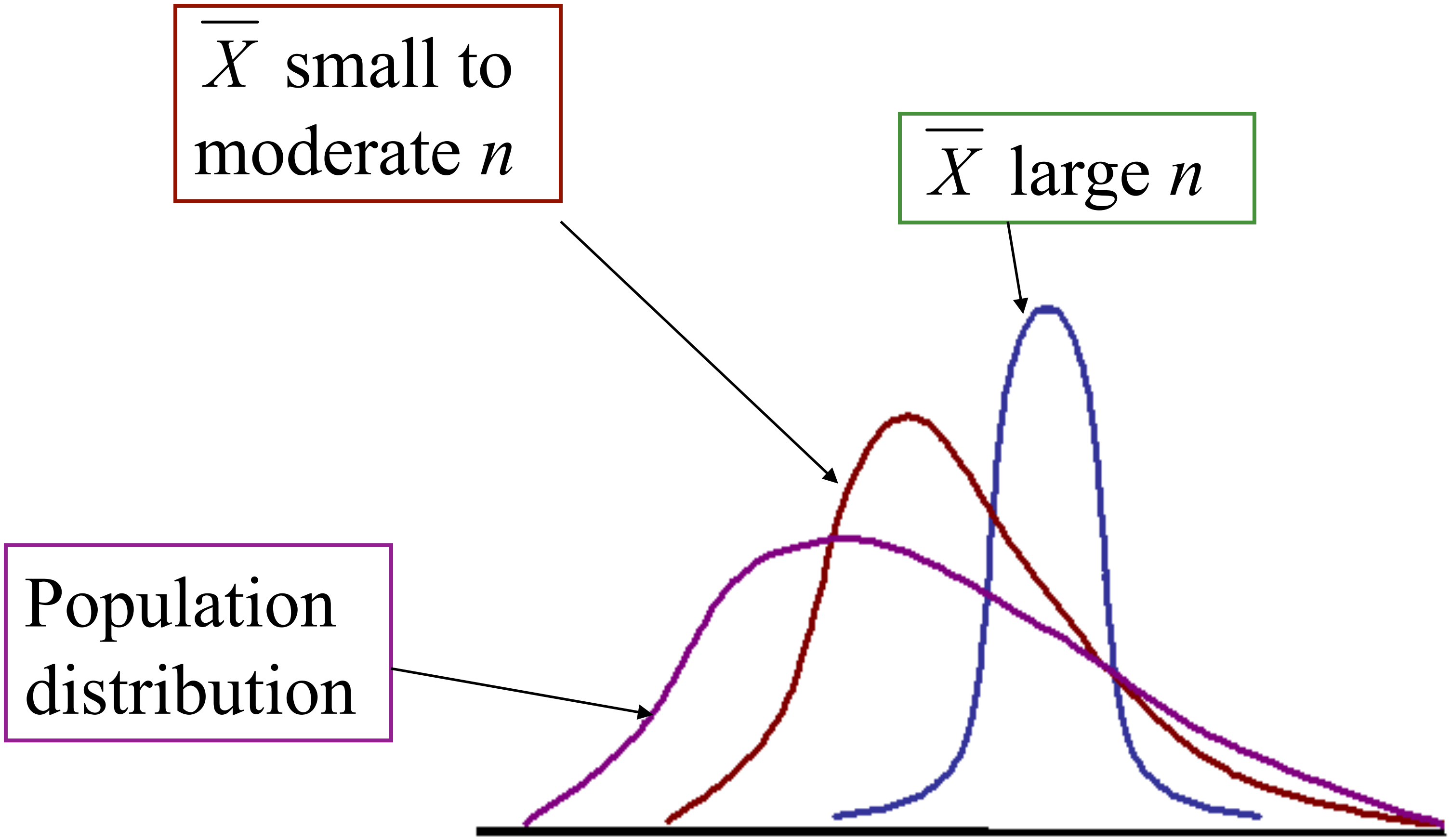
Normal Population Distribution

Let $X_1, \ldots, X_n$ be a random sample from a normal distribution with mean value μ and standard deviation σ. Then for any n, $\overline{X}$ is normally distributed, as is T_o.

The Central Limit Theorem

Let $X_1, \ldots, X_n$ be a random sample from a distribution with mean value μ and variance σ^2. Then if n sufficiently large, $\overline{X}$ has approximately a normal distribution with $\mu_{\overline{X}} = \mu$ and $\sigma^2_{\overline{X}} = \sigma^2 / n$, and T_o also has approximately a normal distribution with $\mu_{T_o} = n\mu$, $\sigma_{T_o} = n\sigma^2$. The larger the value of n, the better the approximation.

The Central Limit Theorem



Rule of Thumb

If $n > 30$, the Central Limit Theorem can be used.

Example/Theorem: Suppose $X_1, \ldots, X_n \overset{\text{iid}}{\sim} f(x)$ with $\mathsf{E}[X_i] = \mu$ and $\mathsf{Var}(X_i) = \sigma^2$. Define the **sample mean** as

$$\bar{X} \equiv \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Then

$$\mathsf{E}[\bar{X}] = \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} \mathsf{E}[X_i] = \frac{1}{n}\sum_{i=1}^{n} \mu = \mu.$$

So the mean of $\bar{X}$ is the same as the mean of X_i.

Meanwhile,. . .

$$\begin{aligned}\text{Var}(\bar{X}) &= \text{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)\\ &= \frac{1}{n^2}\text{Var}\left(\sum_{i=1}^{n} X_i\right)\\ &= \frac{1}{n^2}\sum_{i=1}^{n}\text{Var}(X_i) \quad (X_i\text{'s indep})\\ &= \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \sigma^2/n.\end{aligned}$$

So the mean of $\bar{X}$ is the same as the mean of X_i, but the *variance decreases*!