ECO220Y

Introduction to Probability

Readings: Chapter 6 (skip section 6.9) and Chapter 9 (section 9.1-9.3)

Fall 2011

Lecture 6
Part 2

From Contingency Table to Joint Probability Table

	Student	Staff	Total
American	107	105	212
European	33	12	45
Asian	55	47	102
Total	195	164	359

	Student	Staff	Total
American	$107/359$ $=0.298$	$105/359$ $=0.292$	$212/359$ $=0.59$
European	$33/359$ $=0.092$	$12/359$ $=0.03$	$45/359$ $=0.125$
Asian	$55/359$ $=0.153$	$47/359$ $=0.131$	$102/359$ $=0.284$
Total	$195/359$ $=0.543$	$164/359$ $=0.457$	1

Interpretation of Probabilities

- General definition - long-run frequency
- Interpretation depends on the type of probability:
  - Marginal - proportion/fraction in total population of interest
  - Conditional - proportion/fraction in a group within the population
  - Joint - proportion/fraction that has two attributes out of all attributes defined for a population
- Can define probability as a fraction of population who possesses a certain attribute
- Can define probability as a chance that a randomly selected individual/item possesses a certain attribute

From Words to Joint Probability Table

Employment data at a large company reveal that 72% of the workers are married, 44% are college graduates, and half of the college grads are married.

	College graduates	Not college graduates	Total
Married	$\boxed{0.5*0.44=0.22}$	0.50	0.72
Not married	0.22	0.06	0.28
	0.44	0.56	1

What is the probability that a randomly chosen worker is

(A) Neither married nor college graduate?
(B) Married but not a college graduate?
(C) Married or a college graduate?

Random Variable

- A random variable is defined as numerical outcome of a random experiment.
- Random because we do not know the outcome with certainty.
- Variable because outcome will vary when experiment is repeated.

Think of an outcome of a random experiment when we toss fair coin 5 times.

O_1={H H H H H} O_2={T T T T T}

O_3={H H T H H} O_4={T T H T T}

O_5={H T T T H} O_6={T H H H T}

O_7={H T H T H} O_8={T H T H T}

... ...

If we define random variable X = number of heads in five tosses, what will be the value of random variable X for the listed outcomes?

Types of Random Variables

- Discrete RV - takes a countable (finite) number of values.
  - Number of children per household.
  - Sum of two dice.
- Continuous RV - takes an uncountable (infinite) number of values.
  - Time at checkout in minutes.
  - **Average** number of children for a random sample of households.

Which of the following is a discrete random variable?

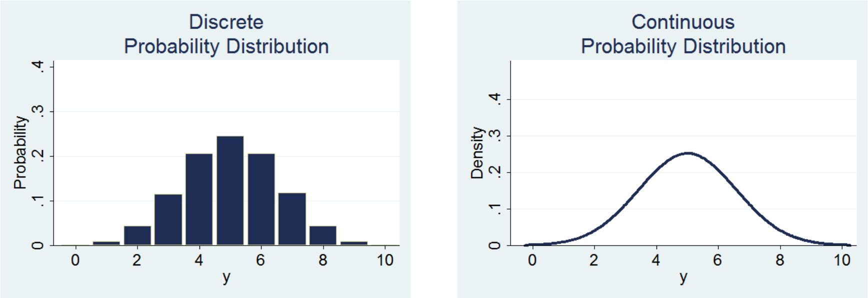
(I) The average height of a randomly selected group of 6-graders.
(II) The annual number of Lotto MAX winners from Toronto.
(III)The number of presidential elections in the 20th century.

Answers:
(A) I only
(B) II only
(C) III only
(D) I and II
(E) II and III

Which of the following is a discrete random variable?

(I) The average height of a randomly selected group of 6-graders.
(II) The annual number of Lotto MAX winners from Toronto.
(III)The number of presidential elections in the 20th century.

Answers:
(A) I only
(B) II only
(C) III only
(D) I and II
(E) II and III



What is the difference between two graphs?

1. Bar chart vs Histogram
2. Probability vs Density
3. Area under the curve is equal to 1 (why?)
4. Sum of the bar heights is equal to 1

Random Variable (Example)

Consider a random experiment: toss a fair coin 3 times, record the number of heads, define X as a number of heads. List all possible outcomes:

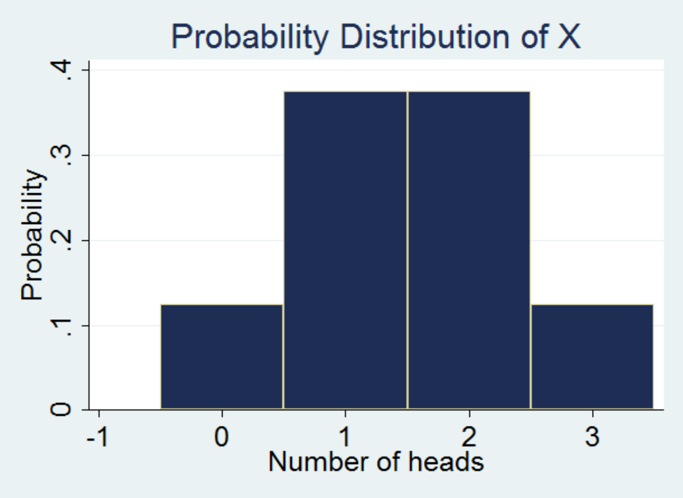
outcome	HHH	HHT	HTT	HTH	THT	TTH	THH	TTT
X=x	3	2	1	2	1	1	2	0

What is the relative frequency of each value? What is the probability of each outcome?

Let's list all the unique values of X with their respective probabilities:

x	P(X=x)
0	1*1/8=1/8
1	3*1/8=3/8
2	3*1/8=3/8
3	1*1/8=1/8
Total	1

We will denote probability that a random variable **X** takes value **x** as P(X=x), or simply p(x).



Mean of A Random Variable

- Probabilities can be interpreted as long-run frequencies.
- For instance, if the probability of heads is .5, then heads should come up about half the time in a large number of flips.
- A similar logic applies to the number of heads X obtained in three coin tosses.
- We already know that probabilities are P(0)=1/8, P(1)=3/8, P(2)=3/8, and P(3)=1/8.
- If we were to toss three coins many, many times, we could anticipate obtaining X=0 and X=3 about one-eighth of the time, X=1 and X=2 three-eighth of the time.
- The average value of X would be 1.5.
- This predicted long-run average is called the expected value.

Expected Value

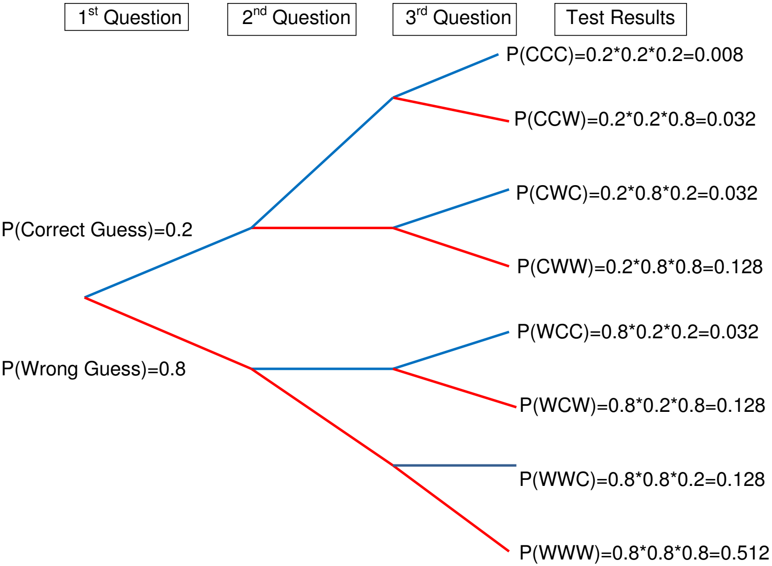
The **expected value** (or mean) of a random variable X is defined as

$$\mu = E[X] = \sum x * p(x)$$

The expected value is a weighted average of the possible outcomes, with the probability weights reflecting how likely each outcome is.

Example - Expected Value of Guess

X is the number of correctly guessed (uninformed) answers to 3 multiple choice questions, each with 5 alternatives.


1st Question
2nd Question
3rd Question
Test Results
P(CCC)=0.2*0.2*0.2=0.008
P(CCW)=0.2*0.2*0.8=0.032
P(CWC)=0.2*0.8*0.2=0.032
P(CWW)=0.2*0.8*0.8=0.128
P(WCC)=0.8*0.2*0.2=0.032
P(WCW)=0.8*0.2*0.8=0.128
P(WWC)=0.8*0.8*0.2=0.128
P(WWW)=0.8*0.8*0.8=0.512
P(Correct Guess)=0.2
P(Wrong Guess)=0.8

Example Cont'd

E[X]=$\sum x * p(x)$=

0*0.512+

1*0.384+

2*0.096+

3*0.008=0.6

Interpretation?

x	p(x)
0	0.512
1	0.384
2	0.096
3	.008

Is X discrete or continuous?

Example Cont'd

Why do you think the following calculation is wrong?

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} = \frac{\sum_{i=1}^{4} x_i}{4} = \frac{0+1+2+3}{4} = 1.5$$

Variance of Random Variable

The **variance of a random variable** X is defined as

$$\sigma^2 = E[(X-\mu)^2] = \sum (x-\mu)^2 * p(x)$$

The **standard deviation** σ is the square root of the variance.

Interpretation: The standard deviation measures the dispersion in the possible outcomes and graphically gauges how "spread out" the distribution is.

If we expect loaded die always come up 6, then the expected value and variance are 6 and 0 respectively. Variance of 0 means that there is no uncertainty about the outcome.

Example Cont'd

$$V[X] = \sum (x - \mu)^2 * p(x) =$$

$$(0 - 0.6)^2 * 0.512 +$$

$$(1 - 0.6)^2 * 0.384 +$$

$$(2 - 0.6)^2 * 0.096 +$$

$$(3 - 0.6)^2 * 0.008 = 0.48$$

x	p(x)
0	0.512
1	0.384
2	0.096
3	.008
μ_x	0.6

Laws of Expected Value and Variance

Laws of Expectation

- E(a)= a
- E(a*X)=a*E(X)
- E(X + a)=E(X) + a

Laws of Variance

- V(a) = 0
- V(X + a)= V(X)
- V(a*X) = a^2*V(X)

Linear Transformation

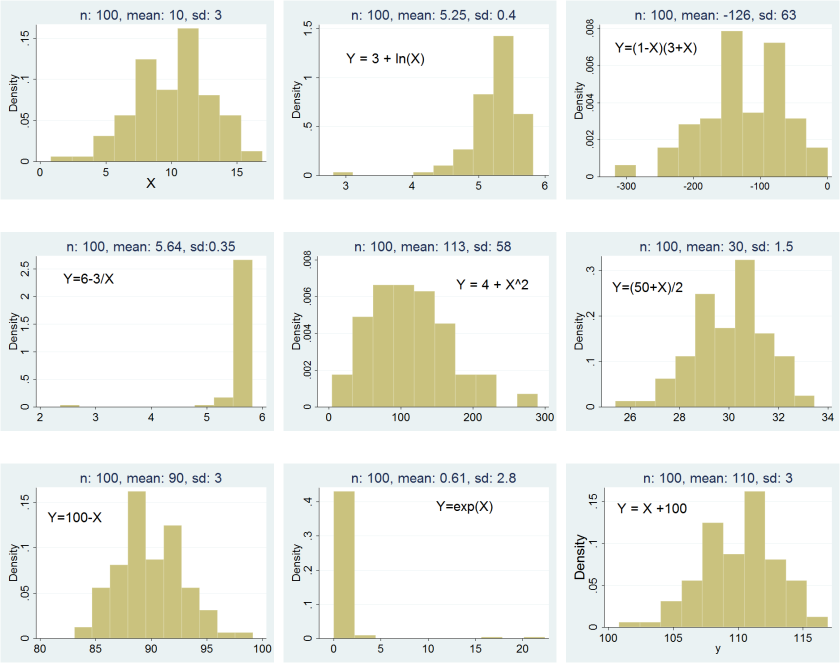
Sometimes, our analysis will require us to look at random variables that are linear transformation of other random variables.

For instance, if we need quickly convert the average temperature in May in Ontario measured in degrees of Celsius (random variable X), then 32+1.8*X is the temperature in degrees Fahrenheit.

What about the <u>shape</u> of the distribution?

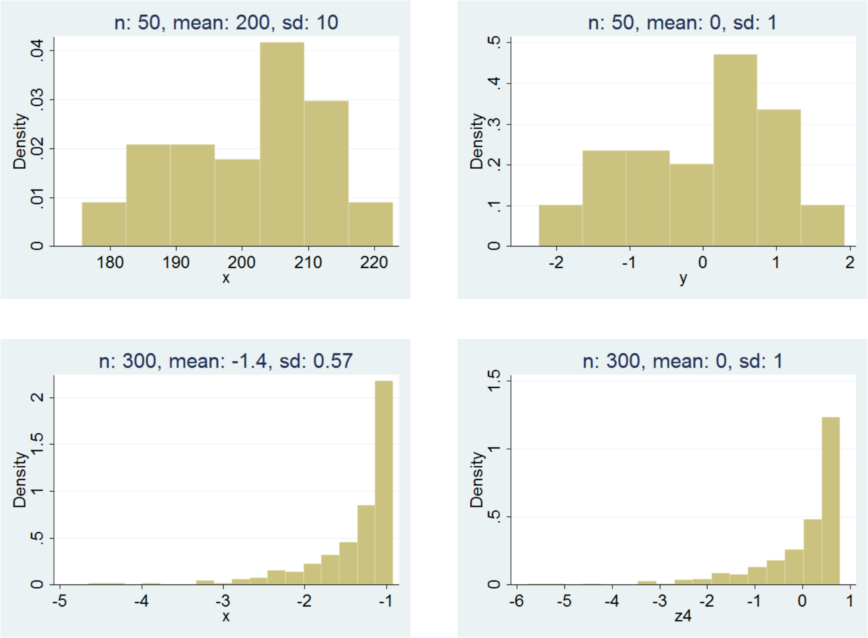
Linear Transformation

Linear Transformation	Linear Transformation?
can be written as	Y=100 - X
$Y = a + b * X$	Y=(50 + X)/2
where a and b are constants.	Y=3 + ln(X)
Holds for both discrete and	Y=6 - 3/X
continuous variables.	Y=exp(X)
Can apply laws of	Y=(1 - X)(3 + X)
expectation	Y=4 + X^2
and variance.	Y=X + π


n: 100, mean: 10, sd: 3
Density
X
n: 100, mean: 5.25, sd: 0.4
Y = 3 + ln(X)
Density
n: 100, mean: -126, sd: 63
Y=(1-X)(3+X)
Density
n: 100, mean: 5.64, sd:0.35
Y=6-3/X
Density
n: 100, mean: 113, sd: 58
Y = 4 + X^2
Density
n: 100, mean: 30, sd: 1.5
Y=(50+X)/2
Density
n: 100, mean: 90, sd: 3
Y=100-X
Density
n: 100, mean: 0.61, sd: 2.8
Y=exp(X)
Density
n: 100, mean: 110, sd: 3
Y = X +100
Density
y

Z-score

- We can standardize random variable.
- Standardize means to apply linear transformation such that the resulting variable has mean 0 and standard deviation equal to 1.
- We call it z-score.
- Z-score tells us how many standard deviation a data point is from the mean.
- Next time, instead of your absolute score on a quiz, I will report your z-score together with the class mean and standard deviation. How would you find your absolute score?


n: 50, mean: 200, sd: 10
Density
0
.01
.02
.03
.04
180
190
200
210
220
x
n: 50, mean: 0, sd: 1
Density
0
.1
.2
.3
.4
.5
-2
-1
0
1
2
y
n: 300, mean: -1.4, sd: 0.57
Density
0
.5
1
1.5
2
-5
-4
-3
-2
-1
x
n: 300, mean: 0, sd: 1
Density
0
.5
1
1.5
-6
-5
-4
-3
-2
-1
0
1
z4

Z-score

$$Z = \frac{X - E[X]}{\sqrt{V[X]}} \text{ or } z = \frac{x - \mu_x}{\sigma_x}$$

Let E[X]=100, and V[X]=25

If Y=(X-100)/5= -20+1/5*X, then what is:

E[Y]= ?

V[Y]= ?

To find it, we need to apply Laws of Expectation and Variance.

Linear Transformation - Example

The average salary at Monster Inc. is \$30,000 a year and variance is equal to \$4,000,000. This year, management awards the following bonuses to each employee:

1. A Christmas bonus of \$500.
2. An incentive bonus equal to 10% of employee's salary.

What is the mean bonus and standard deviation of bonuses received by the employees of Monster Inc.? How much does Monster Inc need to allocate on employees' bonuses this year if it has 40 workers?

- Let's find the mean bonus of employees at Monster Inc
- Let's denote salary X with E[X]=$30,000 and V[X]=$4,000,000
- Bonus (B) is a linear transformation of X.
- E[B]=E[$500+0.1*X] and V[B]=V[$500+0.1*X]
- Let's use Laws of Expected Value and Variance
- E[B]=E[$500]+0.1*E[X]=$500+0.10*$30,000=$3,500
- V[B]=V[$500]+$(0.1)^2$*V[X]=0.01*$4,000,000=$40,000
- SD[B]=$\sqrt{V[B]}$=$200
- Total[B]=E[40*B]=40*E[B]=40*3,500=140,000

Univariate and Bivariate Distributions

1. Univariate Probability Distribution: gives probability of each possible value for a single random variable.
2. Bivariate (or Joint) Probability Distribution: gives probability of each possible pair of values for two random variables.

Can you think of examples of bivariate distribution?

Example

The researcher observes customers in a fast-food restaurant.

X= # of soft drinks ordered

Y= # of sandwiches ordered

Are X and Y interval variables?

What about relationship

between X and Y?

		Sandwich	
		1	2
Soft Drink	1	0.63	0.18
	2	0.09	0.10

Bivariate Distributions

1. Discrete random variables: joint probability is P(X=x and Y=y), or p(x,y). Sum of probabilities over x and y must equal 1.

2. Continuous random variables: joint density is equal to f(X=x and Y=y) or f(x, y). Probability is the volume under the joint density function: volume must be equal to 1.

Covariance and Correlation

$$\text{Covariance: } \sigma_{xy} = E[(X-\mu_x)(Y-\mu_y)]$$

$$E[(x-\mu_x)(Y-\mu_y)] = \sum_{\{all\ x\}} \sum_{\{all\ y\}} (x-\mu_x)(y-\mu_y) * p(x,y)$$

$$\text{Correlation: } \rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

Example Cont'd

	Sandwich		
		1	2	
Soft Drink	1	0.63	0.18	0.81
	2	0.09	0.10	0.19
		0.72	0.28	1.00

First, need to find marginal probabilities:

P(X=1)=P(X=1 and Y=1)+P(X=1 and Y=2)=0.63+0.18=0.81

P(X=2)=P(X=2 and Y=1)+P(X=2 and Y=2)=0.09+0.10=0.19

P(Y=1)=P(Y=1 and X=1)+P(Y=1 and X=2)=0.63+0.09=0.72

P(Y=2)=P(Y=2 and X=1)+P(Y=2 and X=2)=0.18+0.10=0.28

Example Cont'd

	Sandwich			
		1	2	
Soft Drink	1	0.63	0.18	0.81
	2	0.09	0.10	0.19
		0.72	0.28	1.00

$$\mu_X = 1 * 0.81 + 2 * 0.19 = 1.19$$

$$\mu_Y = 1 * 0.72 + 2 * 0.28 = 1.28$$

$$\sigma_X^2 = (1 - 1.19)^2 * 0.81 + (2 - 1.19)^2 * 0.19 = 0.1539, \sigma_X = 0.3923$$

$$\sigma_Y^2 = (1 - 1.28)^2 * 0.72 + (2 - 1.28)^2 * 0.28 = 0.2016, \sigma_Y = 0.4489$$

$$\sigma_{XY} = (1 - 1.19)(1 - 1.28) * 0.63 + (1 - 1.19)(2 - 1.28) * 0.18$$

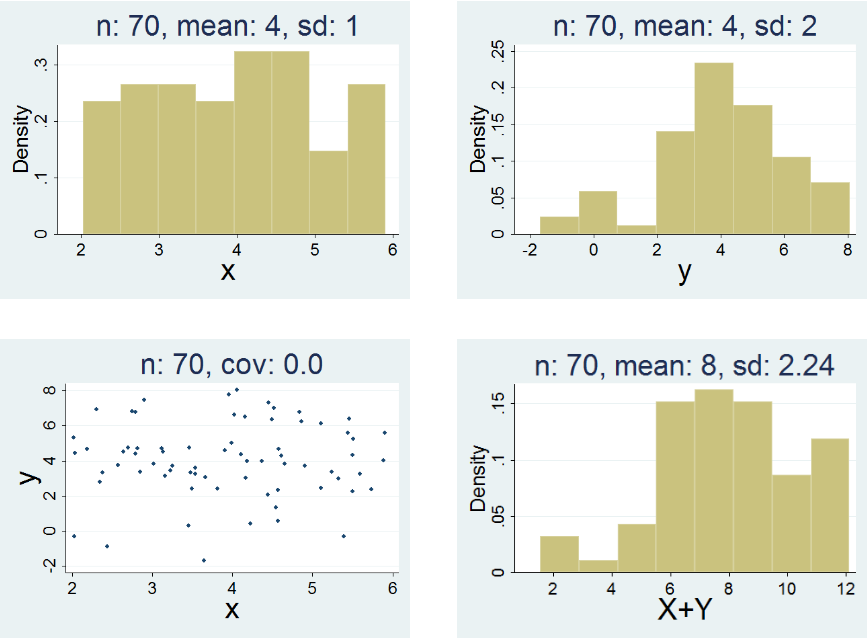
$$+(2 - 1.19)(1 - 1.28) * 0.09 + (2 - 1.19)(2 - 1.28) * 0.10 = 0.0468$$

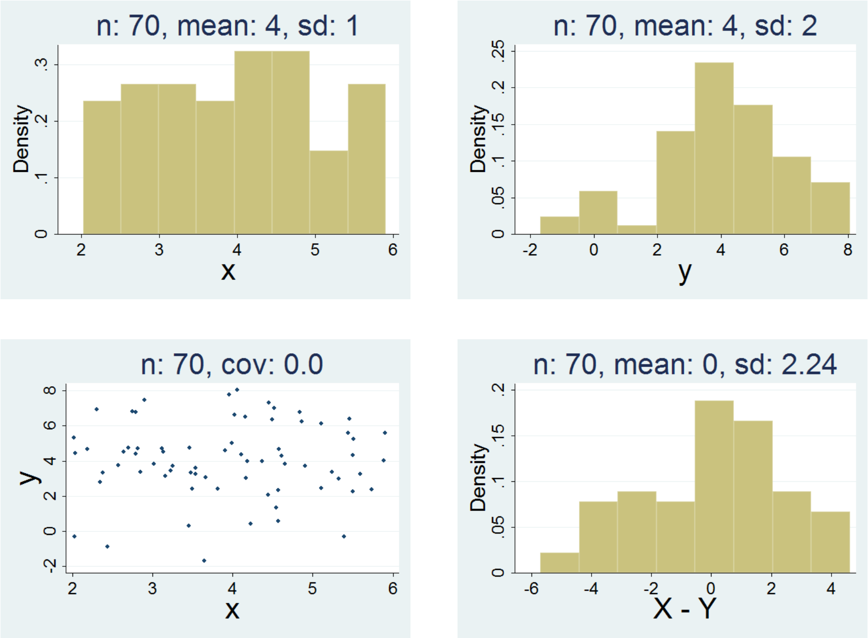
Extending Laws of Expected Value

- $E[a] = a$
- $E[X + a] = E[X] + E[a] = E[X] + a$
- $E[aX] = aE[X]$
- $E[X_1 + X_2 + ... + X_n] = E[X_1] + E[X_2] + ... + E[X_n]$
- $E[a + bX + cY] = E[a + bX] + E[cY] = E[a] + bE[X] + cE[Y] = a + bE[X] + cE[Y]$

Laws of Variance and Covariance

- $COV(X, c) = 0$ (Example? Intuition?)
- $COV(X + a, Y + b) = COV(X, Y)$
- $COV(aX, bY) = a * b * COV(X, Y)$
- $V(c) = 0$
- $V(X + c) = V(X)$
- $V(cX) = c^2 V(X)$
- $V(aX + bY) = a^2 V(X) + b^2 V(Y) + 2abCOV(X, Y)$
- $V(aX + bY) = a^2 V(X) + b^2 V(Y) + 2ab * r(X, Y) * SD(X) * SD(Y)$
- What is V(X-Y) ?


n: 70, mean: 4, sd: 1
Density
x
n: 70, mean: 4, sd: 2
Density
y
n: 70, cov: 0.0
y
x
n: 70, mean: 8, sd: 2.24
Density
X+Y


n: 70, mean: 4, sd: 1
Density
x
n: 70, mean: 4, sd: 2
Density
y
n: 70, cov: 0.0
y
x
n: 70, mean: 0, sd: 2.24
Density
X - Y

Risk and Return

Consider the following four alternatives:

1. \$5 with certainty
2. \$10 with probability 0.5 and \$0 with probability 0.5
3. \$5 with probability 0.5, \$10 with probability 0.25 and \$0 with probability 0.25
4. \$5 with probability 0.5, \$105 with probability 0.25 and -\$95 with probability 0.25

Which alternative will risk-averse individual choose? Risk-seeker?

Compare mean and variance:

1. Expected return=\$5, standard deviation=\$0
2. Expected return=\$5, standard deviation=\$5
3. Expected return=\$5, standard deviation=\$3.54
4. Expected return=\$5, standard deviation=\$70.71

Application: Risk and Return

If we know the rates of return and associated risk for each of the following stocks:

Stock	RV	Return E[RV]	Riskiness V[RV]	Share	Correlation r
McDonald's	X	8%	10%	60%	-0.3
Motorola	Y	3%	6%	40%	-0.3

we can find the return on our portfolio as well as how risky is our portfolio.

Application: Risk and Return

Let W be a random variable such that W=0.6X+0.4Y

Then,
E[W]=E[0.6X+0.4Y]=0.6*E[X]+0.4*E[Y]=0.6*0.08+0.4*0.03=0.06, or 6%

V[W]=V[0.6X+0.4Y]=0.6^2*V[X]+0.4^2*V[Y]=0.6^2*0.10+0.4^2*0.06=0.0456

Is this everything?

V[W]=0.6^2*V[X]+0.4^2*V[Y]+2*0.6*0.4*(-0.3)*$\sqrt{0.10 * 0.06}$=0.0345

Risk and Return Again

State of Economy	Probability	E[X]	E[Y]	E[W]=0.5X+0.5Y
Recession	25%	2%	2%	2%
Moderate Growth	50%	8%	8%	8%
Boom	25%	14%	14%	14%
Expected Return		8%	8%	8%
St.Deviation		4.24%	4.24%	4.24%

Correlation	1

Risk and Return Again

State of Economy	Probability	E[X]	E[Y]	E[W]=0.5X+0.5Y
Recession	25%	2%	14%	8%
Moderate Growth	50%	8%	8%	8%
Boom	25%	14%	2%	8%
Expected Return		8%	8%	8%
St.Deviation		4.24%	4.24%	0%

Correlation -1

Risk and Return Again

State of Economy	Probability	E[X]	E[Y]	E[W]=0.5X+0.5Y
Recession	25%	2%	2%	2%
Moderate Growth	50%	8%	2%	5%
Boom	25%	14%	2%	8%
Expected Return		8%	2%	5%
St.Deviation		4.24%	0%	2.12%

Correlation 0

Conclusion: The lower the correlation, the more risk reduction