Math C067 — Solutions

Richard Beigel

Last revision date: some time between March 28, 2007 and April 18, 2007

Bring homework to class with you. Don't have someone else hand it in for you. Don't even hand it in yourself. We will go over homework at the start of each lecture. Please be there.

1. Preliminaries

   1-1. Buy the textbook. If this means paying Amazon to fedex it to you, pay. It's your education, and it's also 1/2 a point towards your final grade.

   1-2. If you don't have a calculator yet, buy one. My students seem to prefer statistical calculators, but any calculator is ok as long as you are comfortable using it.

2. Descriptive Statistics

   2-1. Read the Portfolio Selection handout. Identify 6 statistical terms in the Portfolio Selection handout and find the formulas for calculating them in Chapter 1 of your textbook. Mark those formulas with post-it tabs.

   - variance: page 8, formulas (1.3) and (1.5)
   - weighted average: page 15, first boxed formula (it's called ıweighted mean in the textbook, so it's ok if you didn't find this one)
   - mean: page 5, formula (1.1)
   - standard deviation: page 8, formula (1.4)
   - covariance: page 19, formula (1.10)
   - correlation: page 19, formulas (1.9), (1.8), and (1.11)

   2-2. We are given the following data: $x_1 = 2$, $x_2 = 3$, $x_3 = 5$, and $x_4 = 7$. Calculate $\sum_{i=1}^{4} x_i$.

$$2 + 3 + 5 + 7 = 17$$

   2-3. Calculate $\sum_{i=1}^{4} i$.

$$1 + 2 + 3 + 4 = 10$$

   2-4. Research on X Corportation Stock turned up the following returns for 4 different years: \$0, \$0, \$8, and \$0.

   (a) What is the mean of those 4 returns?

$$\bar{x} = (0 + 0 + 8 + 0)/4 = 2$$

   (b) What is the variance of those returns?

$$\begin{aligned} s^2 &= \left((0-2)^2 + (0-2)^2 + (8-2)^2 + (0-2)^2\right)/3 \\ &= \left((-2)^2 + (-2)^2 + 6^2 + (-2)^2\right)/3 \\ &= (4 + 4 + 36 + 4)/3 \\ &= 48/3 \\ &= 16 \end{aligned}$$

(c) What is the standard deviation of those returns?

$$s = \sqrt{s^2} = \sqrt{16} = 4$$

(d) What is the median of those returns?

$$(0 + 0)/2 = 0$$

2-5. The returns on Y Corportation Stock in the same 4 years were –\$1, \$4, \$3, and \$2.

(a) What is the mean of those 4 returns?

$$\bar{y} = (-1 + 4 + 3 + 2)/4 = 2$$

(b) What is the variance of those returns?

$$\begin{aligned} s^2 &= \left((-1-2)^2 + (4-2)^2 + (3-2)^2 + (2-2)^2\right)/3 \\ &= \left((-3)^2 + 2^2 + 1^2 + 0^2\right)/3 \\ &= (9 + 4 + 1 + 0)/3 \\ &= 14/3 \\ &\approx 4.67 \end{aligned}$$

(c) What is the standard deviation of those returns?

$$s = \sqrt{s^2} = \sqrt{14/3} \approx \sqrt{4.67} \approx 2.16$$

(d) What is the median of those returns?

$$(2 + 3)/2 = 2.5$$

(e) Which investment would you prefer, X or Y? Why? Both stocks have the same average return. Because Stock Y has lower variance, Stock Y is less risky, so I would prefer Stock Y.

2-6. What is the covariance between the Stock X and Stock Y returns (refer to previous homework)? Recall that $\bar{x} = 2$ and $\bar{y} = 2$.

$$\begin{aligned} s_{xy} &= \left((x_1 - \bar{x})(y_1 - \bar{y}) + \cdots + (x_n - \bar{x})(y_n - \bar{y})\right)/(n-1) \\ &= \left((0-2)(-1-2) + (0-2)(4-2) + (8-2)(3-2) + (0-2)(2-2)\right)/3 \\ &= \left((-2)(-3) + (-2)(2) + 6(1) + (-2)(0)\right)/3 \\ &= (6 + -4 + 6 + 0)/3 \\ &= 8/3 \\ &\approx 2.67 \end{aligned}$$

2-7. What is the correlation between the Stock X and Stock Y returns? Recall that $s_x = 4$ and $s_y \approx 2.16$.

$$r = \frac{s_{xy}}{s_x s_y} \approx \frac{2.67}{4 \cdot 2.16} \approx 0.31$$

2-8. Determine the regression line of Stock Y vs. Stock X. The best-fit line is $y = a + bx$ where

- $b = r s_y / s_x \approx 0.31 \cdot 2.16/4 \approx 0.17$ and
- $a = \bar{y} - b\bar{x} \approx 2 - 0.17 \cdot 2 = 1.66$

2-9. (Scatter)Plot the Stock X and Stock Y returns as (X,Y) pairs. Graph the regression line as well. Does the regression line seem to fit the data points? If not, re-calculate the regression line.

2-10. Consider a Mutual Fund Z that invests 1/3 of its money in Stock X and 2/3 in Stock Y. Its returns are given by the formula

$$z_i = \frac{1}{3}x_i + \frac{2}{3}y_i$$

(a) Calculate the returns on Mutual Fund Z.

- $z_1 = \frac{1}{3}x_1 + \frac{2}{3}y_1 = \frac{1}{3}\cdot 0 + \frac{2}{3}\cdot(-1) = 0 + -\frac{2}{3} \approx -0.667$
- $z_2 = \frac{1}{3}x_2 + \frac{2}{3}y_2 = \frac{1}{3}\cdot 0 + \frac{2}{3}\cdot 4 = 0 + \frac{8}{3} \approx 2.667$
- $z_3 = \frac{1}{3}x_3 + \frac{2}{3}y_3 = \frac{1}{3}\cdot 8 + \frac{2}{3}\cdot 3 = \frac{8}{3} + 2 \approx 4.667$
- $z_4 = \frac{1}{3}x_4 + \frac{2}{3}y_4 = \frac{1}{3}\cdot 0 + \frac{2}{3}\cdot 2 = 0 + \frac{4}{3} \approx 1.333$

(b) What is the fund's average return?

$$\bar{z} \approx (-0.667 + 2.667 + 4.667 + 1.333)/4 = 2$$

Alternative solution:

$$\bar{z} = \frac{1}{3}\bar{x} + \frac{2}{3}\bar{y} = \frac{1}{3}\cdot 2 + \frac{2}{3}\cdot 2 = 2$$

(c) What is its variance?

$$\begin{aligned} s^2 &\approx \left((-0.667-2)^2 + (2.667-2)^2 + (4.667-2)^2 + (1.333-2)^2\right)/3 \\ &= \left((-2.667)^2 + 0.667^2 + 2.667^2 + (-0.667)^2\right)/3 \\ &\approx 15.1/3 \\ &\approx 5.04 \end{aligned}$$

(d) Would you prefer Stock X, Stock Y, or Mutual Fund Z? All three choices have the same expected return (mean). Recall that the variance of Stock X is 16 and the variance of Stock Y is approximately 4.67. Since Stock Y has the lowest risk (variance), Stock Y is preferable.

2-11. Consider a Mutual Fund M that invests 10% of its money in Stock X and 90% in Stock Y.

(a) Give a formula for the returns on M in each year. If you can't do this part, ask for help so you can do the remaining parts.

$$m_i = \frac{1}{10}x_i + \frac{9}{10}y_i$$

(b) Calculate the returns on Mutual Fund M.

- $m_1 = \frac{1}{10}x_1 + \frac{9}{10}y_1 = \frac{1}{10}\cdot 0 + \frac{9}{10}\cdot(-1) = 0 + -\frac{9}{10} = -0.9$
- $m_2 = \frac{1}{10}x_2 + \frac{9}{10}y_2 = \frac{1}{10}\cdot 0 + \frac{9}{10}\cdot 4 = 0 + \frac{36}{10} = 3.6$
- $m_3 = \frac{1}{10}x_3 + \frac{9}{10}y_3 = \frac{1}{10}\cdot 8 + \frac{9}{10}\cdot 3 = \frac{8}{10} + \frac{27}{10} = \frac{35}{10} = 3.5$
- $m_4 = \frac{1}{10}x_4 + \frac{9}{10}y_4 = \frac{1}{10}\cdot 0 + \frac{9}{10}\cdot 2 = 0 + \frac{18}{10} = 1.8$

(c) What is the fund's average return?

$$\bar{m} = (-0.9 + 3.6 + 3.5 + 1.8)/4 = 2$$

Alternative solution:

$$\bar{m} = \frac{1}{10}\bar{x} + \frac{9}{10}\bar{y} = \frac{1}{10}\cdot 2 + \frac{9}{10}\cdot 2 = 2$$

(d) What is its variance?

$$\begin{aligned} s^2 &\approx \left((-0.9-2)^2 + (3.6-2)^2 + (3.5-2)^2 + (1.8-2)^2\right)/3 \\ &= \left((-2.9)^2 + 1.6^2 + 1.5^2 + (-0.2)^2\right)/3 \\ &= 13.26/3 \\ &= 4.42 \end{aligned}$$

(e) Would you prefer Stock X, Stock Y, Mutual Fund Z, or Mutual Fund M? All four choices have the same expected return (mean). Recall that the variance of Stock X is 16 and the variance of Stock Y is approximately 4.67. Since Mutual Fund M has the lowest risk (variance), Fund M is preferable.

2-12. Use the following data for Stock X and Stock Y.

Year	1	2	3	4	5
Return on Stock X	1	3	5	7	9
Return on Stock Y	7	6	5	4	3

A mutual fund Z wishes to put a fraction of its money into Stock X and the rest into Stock Y.

(a) How should Fund Z weight Stock X and Stock Y to minimize risk? Use that weighting in the remaining parts of this problem.

- $\bar{x} = (1+3+5+7+9)/5 = 5$. (Note that because the numbers are evenly spaced the mean has to be equal to the middle number.)
-
$$\begin{aligned} s_x^2 &= ((x_1-\bar{x})^2+(x_2-\bar{x})^2+(x_3-\bar{x})^2+(x_4-\bar{x})^2+(x_5-\bar{x})^2)/4 \\ &= ((1-5)^2+(3-5)^2+(5-5)^2+(7-5)^2+(9-5)^2)/4 \\ &= ((-4)^2+(-2)^2+0^2+2^2+4^2)/4 \\ &= (16+4+0+4+16)/4 \\ &= 40/4 \\ &= 10 \end{aligned}$$
- $\bar{y} = (7+6+5+4+3)/5 = 5$. (Note that because the numbers are evenly spaced the mean has to be equal to the middle number.)
-
$$\begin{aligned} s_y^2 &= ((y_1-\bar{y})^2+(y_2-\bar{y})^2+(y_3-\bar{y})^2+(y_4-\bar{y})^2+(y_5-\bar{y})^2)/4 \\ &= ((7-5)^2+(6-5)^2+(5-5)^2+(4-5)^2+(3-5)^2)/4 \\ &= (2^2+1^2+0^2+(-1)^2+(-2)^2)/4 \\ &= (4+1+0+1+4)/4 \\ &= 10/4 \\ &= 2.5 \end{aligned}$$
-
$$\begin{aligned} s_{xy} &= ((x_1-\bar{x})(y_1-\bar{y})+(x_2-\bar{x})(y_2-\bar{y})+(x_3-\bar{x})(y_3-\bar{y})+(4_1-\bar{x})(y_4-\bar{y})+(x_5-\bar{x})(y_5-\bar{y}))/4 \\ &= ((-4)2+(-2)1+0\cdot 0+2(-1)+4(-2))/4 \\ &= (-8+-2+0+-2+-8)/4 \\ &= -20/4 \\ &= -5 \end{aligned}$$
- Let $z_i = wx_i + (1-w)y_i$
$$\begin{aligned} s_z^2 &= w^2 s_x^2 + 2w(1-w)s_{xy} + (1-w)^2 s_y^2 \\ &= w^2\cdot 10 + 2w(1-w)\cdot(-5) + (1-w)^2\cdot 2.5 \\ &= 10w^2 + 2\cdot(-5)w(1-w) + 2.5(1-w)^2 \\ &= 10w^2 + -10(w-w^2) + 2.5(1-2w+w^2) \\ &= 10w^2 - 10w + 10w^2 + 2.5 - 5w + 2.5w^2 \\ &= 10w^2 + 10w^2 + 2.5w^2 - 10w - 5w + 2.5 \\ &= 22.5w^2 - 15w + 2.5 \end{aligned}$$

That quadradratic is minimized when $w = 15/(2\cdot 22.5) = 1/3$

(b) What is the variance of Fund Z for the years studied?

$$\begin{aligned} s_z^2 &= 22.5w^2 - 15w + 2.5 \\ &= 22.5(1/3)^2 - 15(1/3) + 2.5 \\ &= 2.5 - 5 + 2.5 \\ &= 0 \end{aligned}$$

(c) What is the mean of Fund Z for the years studied?

$$\begin{aligned} \bar{z} &= w\bar{x} + (1-w)\bar{y} \\ &= (1/3)\cdot 5 + (1-1/3)\cdot 5 \\ &= 5 \end{aligned}$$

2-13. Use the following data for Stock X and Stock Y.

Year	1	2	3	4	5
Return on Stock X	1	3	5	7	9
Return on Stock Y	7	3	4	5	6

A mutual fund Z wishes to put a fraction of its money into Stock X and the rest into Stock Y.

(a) How should Fund Z weight Stock X and Stock Y to minimize risk? Use that weighting in the remaining parts of this problem.

- The Stock X data is the same as in the previous problem, so $\bar{x} = 5$ and $s_x^2 = 10$.
- The Stock Y data is a rearrangement of the Stock Y data from the previous problem, so $\bar{y} = 5$ and $s_y^2 = 2.5$.
- $$\begin{aligned} s_{xy} &= ((x_1-\bar{x})(y_1-\bar{y}) + (x_2-\bar{x})(y_2-\bar{y}) + (x_3-\bar{x})(y_3-\bar{y}) + (4_1-\bar{x})(y_4-\bar{y}) + (x_5-\bar{x})(y_5-\bar{y}))/4 \\ &= ((1-5)(7-5) + (3-5)(3-5) + (5-5)(4-5) + (7-5)(5-5) + (9-5)(6-5))/4 \\ &= ((-4)(2) + (-2)(-2) + (0)(-1) + (2)(0) + (4)(1))/4 \\ &= (-8+4+0+0+4)/4 \\ &= 0/4 \\ &= 0 \end{aligned}$$
- Let $z_i = wx_i + (1-w)y_i$

$$\begin{aligned} s_z^2 &= w^2 s_x^2 + 2w(1-w)s_{xy} + (1-w)^2 s_y^2 \\ &= w^2\cdot 10 + 2w(1-w)\cdot(0) + (1-w)^2\cdot 2.5 \\ &= 10w^2 + 2.5(1-w)^2 \\ &= 10w^2 + 2.5(1-2w+w^2) \\ &= 10w^2 + 2.5 - 5w + 2.5w^2 \\ &= 12.5w^2 - 5w + 2.5 \end{aligned}$$

That quadradratic is minimized when $w = 5/(2\cdot 12.5) = 1/5$

(b) What is the variance of Fund Z for the years studied?

$$\begin{aligned} s_z^2 &= 12.5w^2 - 5w + 2.5 \\ &= 12.5(1/5)^2 - 5(1/5) + 2.5 \\ &= 12.5/25 - 5(1/5) + 2.5 \\ &= 0.5 - 1 + 2.5 \\ &= 2 \end{aligned}$$

(c) What is the mean of Fund Z for the years studied?

$$\begin{aligned}\bar{z} &= w\bar{x} + (1-w)\bar{y} \\ &= (1/5)\cdot 5 + (1-1/5)\cdot 5 \\ &= 5\end{aligned}$$

2-14. Use the following data for Stock X and Stock Y. It's not the same.

Year	1	2	3	4	5
Return on Stock X	1	3	5	7	9
Return on Stock Y	3	4	5	6	7

A mutual fund Z wishes to put a fraction of its money into Stock X and the rest into Stock Y.

(a) How should Fund Z weight Stock X and Stock Y to minimize risk? Use that weighting in the remaining parts of this problem.

- The Stock X data is the same as in the previous problem, so $\bar{x} = 5$ and $s_x^2 = 10$.
- The Stock Y data is a rearrangement of the Stock Y data from the previous problem, so $\bar{y} = 5$ and $s_y^2 = 2.5$.
-

$$\begin{aligned}s_{xy} &= ((x_1-\bar{x})(y_1-\bar{y}) + (x_2-\bar{x})(y_2-\bar{y}) + (x_3-\bar{x})(y_3-\bar{y}) + (4_1-\bar{x})(y_4-\bar{y}) + (x_5-\bar{x})(y_5-\bar{y}))/4 \\ &= ((1-5)(3-5) + (3-5)(4-5) + (5-5)(5-5) + (7-5)(6-5) + (9-5)(7-5))/4 \\ &= ((-4)(-2) + (-2)(-1) + (0)(0) + (2)(1) + (4)(2))/4 \\ &= (8+2+0+2+8)/4 \\ &= 20/4 \\ &= 5\end{aligned}$$

- Let $z_i = wx_i + (1-w)y_i$

$$\begin{aligned}s_z^2 &= w^2 s_x^2 + 2w(1-w)s_{xy} + (1-w)^2 s_y^2 \\ &= w^2\cdot 10 + 2w(1-w)\cdot(5) + (1-w)^2\cdot 2.5 \\ &= 10w^2 + 10w(1-w) + 2.5(1-w)^2 \\ &= 10w^2 + 10(w-w^2) + 2.5(1-w)^2 \\ &= 10w^2 + 10(w-w^2) + 2.5(1-2w+w^2) \\ &= 10w^2 + 10w - 10w^2 + 2.5 - 5w + 2.5w^2 \\ &= 10w^2 - 10w^2 + 2.5w^2 + 10w - 5w + 2.5 \\ &= 2.5w^2 + 5w + 2.5\end{aligned}$$

That quadradratic is minimized when $w = -5/(2\cdot 2.5) = -1$. (Giving a negative weight to a stock means selling it short. This is something that individual investors and hedge funds are allowed to do, but hedge funds are not. A real mutual fund would replace the negative weight by 0 and invest 100% in Stock Y.)

(b) What is the variance of Fund Z for the years studied?

$$\begin{aligned}s_z^2 &= 2.5w^2 + 5w + 2.5 \\ &= 2.5(-1)^2 + 5(-1) + 2.5 \\ &= 2.5 - 5 + 2.5 \\ &= 0\end{aligned}$$

(c) What is the mean of Fund Z for the years studied?

$$\begin{aligned}\bar{z} &= w\bar{x} + (1-w)\bar{y} \\ &= (1/5)\cdot 5 + (1-1/5)\cdot 5 \\ &= 5\end{aligned}$$

2-15. Use the following data for Stock X and Stock Y.

Year	1	2	3	4	5
Return on Stock X	1	3	5	7	9
Return on Stock Y	7	3	4	5	6

A mutual fund Z wishes to put a fraction of its money into Stock X and the rest into Stock Y.

(a) How should Fund Z weight Stock X and Stock Y to minimize risk? Use that weighting in the remaining parts of this problem.

(b) What is the variance of Fund Z for the years studied?

(c) What is the mean of Fund Z for the years studied?

(d) What are returns on Fund Z in years 1 through 5? $z_i = wx_i + (1-w)y_i = 0.2x_i + 0.8y_i$ so

- $z_1 = 0.2x_1 + 0.8y_1 = 0.2\cdot 1 + 0.8\cdot 7 = 5.8$
- $z_2 = 0.2x_2 + 0.8y_2 = 0.2\cdot 3 + 0.8\cdot 3 = 3$
- $z_3 = 0.2x_3 + 0.8y_3 = 0.2\cdot 5 + 0.8\cdot 4 = 4.2$
- $z_4 = 0.2x_4 + 0.8y_4 = 0.2\cdot 7 + 0.8\cdot 5 = 5.4$
- $z_5 = 0.2x_5 + 0.8y_5 = 0.2\cdot 9 + 0.8\cdot 6 = 6.6$

2-16. Use the following data for Stock X and Stock Y. It's not the same.

Year	1	2	3	4	5
Return on Stock X	1	3	5	7	9
Return on Stock Y	3	4	5	6	7

A mutual fund Z wishes to put a fraction of its money into Stock X and the rest into Stock Y.

(a) How should Fund Z weight Stock X and Stock Y to minimize risk? Use that weighting in the remaining parts of this problem.

(b) What is the variance of Fund Z for the years studied?

(c) What is the mean of Fund Z for the years studied?

(d) What are returns on Fund Z in years 1 through 5? $z_i = wx_i + (1-w)y_i = 0.2x_i + 0.8y_i$ so

- $z_1 = (-1)x_1 + 2y_1 = (-1)\cdot 1 + 2\cdot 3 = 5$
- $z_2 = (-1)x_2 + 2y_2 = (-1)\cdot 3 + 2\cdot 4 = 5$
- $z_3 = (-1)x_3 + 2y_3 = (-1)\cdot 5 + 2\cdot 5 = 5$
- $z_4 = (-1)x_4 + 2y_4 = (-1)\cdot 7 + 2\cdot 6 = 5$
- $z_5 = (-1)x_5 + 2y_5 = (-1)\cdot 9 + 2\cdot 7 = 5$

3. Sets and Counting

3-1. Use the following sets for this question:

- $U = \{\mathcal{A}, \ldots, \mathcal{Z}\}$
- $V = \{\mathcal{A}, \mathcal{E}, \mathcal{I}, \mathcal{O}, \mathcal{U}\}$
- $C = \{\mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{F}, \mathcal{G}, \mathcal{H}, \mathcal{J}, \mathcal{K}, \mathcal{L}, \mathcal{M}, \mathcal{N}, \mathcal{P}, \mathcal{Q}, \mathcal{R}, \mathcal{S}, \mathcal{T}, \mathcal{V}, \mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}\}$
- $H = \{\mathcal{H}, \mathcal{P}\}$
- $I = \{\mathcal{I}, \mathcal{B}, \mathcal{M}\}$
- $\phi = \{\}$

(a) What is $C \cap I$? $\{\mathcal{B}, \mathcal{M}\}$

(b) What is $\phi \cap V$? ϕ

(c) What is $V \cup I$? $\{\mathcal{A}, \mathcal{B}, \mathcal{E}, \mathcal{I}, \mathcal{M}, \mathcal{O}, \mathcal{U}\}$

(d) What is $\phi \cup V$? V

(e) $V = \overline{C}$, i.e., $V = C^{c}$. Find two more pairs of sets such that one is the complement of the other.

- $C = \overline{V}$
- $\phi = \overline{U}$
- $U = \overline{\phi}$

3-2. Use the following sets for this question:

- $C = \{\mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{F}, \mathcal{G}, \mathcal{H}, \mathcal{J}, \mathcal{K}, \mathcal{L}, \mathcal{M}, \mathcal{N}, \mathcal{P}, \mathcal{Q}, \mathcal{R}, \mathcal{S}, \mathcal{T}, \mathcal{V}, \mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}\}$
- $H = \{\mathcal{H}, \mathcal{P}\}$
- $I = \{\mathcal{I}, \mathcal{B}, \mathcal{M}\}$
- $\phi = \{\}$

$\phi \subseteq C$. Find four more pairs of sets such that one is a subset of the other.

- $H \subseteq C$
- $\phi \subseteq H$
- $\phi \subseteq I$
- $\phi \subseteq \phi$
- $C \subseteq C$
- $H \subseteq H$
- $I \subseteq I$

3-3. Use the following sets for this question.

- $V = \{\mathcal{A}, \mathcal{E}, \mathcal{I}, \mathcal{O}, \mathcal{U}\}$
- $C = \{\mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{F}, \mathcal{G}, \mathcal{H}, \mathcal{J}, \mathcal{K}, \mathcal{L}, \mathcal{M}, \mathcal{N}, \mathcal{P}, \mathcal{Q}, \mathcal{R}, \mathcal{S}, \mathcal{T}, \mathcal{V}, \mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}\}$
- $H = \{\mathcal{H}, \mathcal{P}\}$
- $I = \{\mathcal{I}, \mathcal{B}, \mathcal{M}\}$
- $\phi = \{\}$

V and C are disjoint. Find all other pairs of disjoint sets. (There are 6 more or 12 more depending on how you count.

- H and I are disjoint
- ϕ and V are disjoint
- ϕ and C are disjoint
- ϕ and H are disjoint
- ϕ and I are disjoint
- ϕ and ϕ are disjoint

3-4. In the town of Oberjoi everyone is rich or beautiful. According to the latest Oberjoi census,

- 100 of the residents are rich
- 200 of the residents are beautiful
- 50 of the residents are rich and beautiful

How many people live in Oberjoi?

- Let A be the set consisting of all rich Oberjoi residents.
- Let B be the set consisting of all beautiful Oberjoi residents.
- We are given that $|A| = 100$, $|B| = 200$, and $|A \cap B| = 50$.
- We are asked to find $|A \cup B|$.

The "Inclusion–Exclusion" principle (formula for the size of a union) says that

$$\begin{aligned} |A \cup B| &= |A| + |B| - |A \cap B| \\ &= 100 + 200 - 50 \end{aligned}$$

3-5. Everyone attending George's birthday party on July 6 was powerful or a Republican

- 80 of the attendees were powerful
- 100 of the attendees were Republicans
- 75 of the attendees were powerful Republicans

How many people attended George's birthday party?

$$80 + 100 - 75 = 105$$

3-6. Everyone attending Jenna and Barbara's birthday party on November 25 got drunk or stoned

- 75 of the partiers got drunk
- 20 of the partiers got stoned
- 5 of the partiers got drunk and stoned

How many people attended Jenna and Barbara's birthday party?

$$75 + 20 - 5 = 90$$

3-7.

$$\begin{aligned} M &= \{\text{chicken, beef, pork, shrimp, plain}\} \\ K &= \{\text{chicken, beef, lamb, tofu}\} \\ E &= \{\text{lo mein, chow mein, egg fu yung}\} \\ D &= \{\text{ice cream, lychee nut}\} \end{aligned}$$

What is $|M \cap K|$?

$$\begin{aligned} M \cap K &= \{\text{chicken, beef}\} \\ |M \cap K| &= 2 \end{aligned}$$

3-8. Use the formula for Cardinality of Union (also called Inclusion Exclusion) to calculate $|M \cup K|$

$$\begin{aligned} |M \cup K| &= |M| + |K| - |M \cap K| \\ &= 5 + 4 - 2 \\ &= 7 \end{aligned}$$

3-9. Use the formula from today's lecture to calculate $|M - K|$

$$\begin{aligned} |M - K| &= |M| - |M \cap K| \\ &= 5 - 2 \\ &= 3 \end{aligned}$$

- Let A be the set consisting of all partiers who were at least 20 years old
- Let B be the set consisting of all partiers who were at most 30 years old
- We are given that $|A| = 100$, $|B| = 150$, and $|A \cup B| = 200$.
- We are asked to find $|A \cap B|$.

The "Inclusion–Exclusion" principle (formula for the size of a union) says that

$$\begin{aligned} |A \cup B| &= |A| + |B| - |A \cap B| \\ 200 &= 100 + 150 - |A \cap B| \\ |A \cap B| &= 100 + 150 - 200 \\ |A \cap B| &= 50 \end{aligned}$$

3-10. Eli's Chinese restaurant serves three different entries

- lo mein, chow mein, and egg fu yung

each prepared with

- chicken, beef, lamb, or tofu.

Eli's menu includes two desserts

- ice cream and lychee nuts

Represent a meal as an ordered triple (meat, entree, dessert). What is the set of possible meals that can be ordered at Eli's? Give your answer in the form $A \times B \times C$ where A, B, and C are sets.

$$K \times E \times D$$

3-11. How many different meals with dessert can be ordered at Eli's?

$$|K \times E \times D| = |K| \cdot |E| \cdot |D| = 4 \cdot 3 \cdot 2 = 24$$

3-12. The Phillies play a 4-game series. The outcome if each game is W, L, or R, where W stands for Win, L stands for Lose, and R stands for rain. Represent a possible outcome for the series as an ordered 4-tuple. What is the set of possible outcomes for the 4-game series? Give your answer in the form A^k where A is a set and k is an integer.

$$\{\mathrm{W}, \mathrm{L}, \mathrm{R}\}^4$$

3-13. How many different outcomes are possible for the 4-game series?

$$|\{\mathrm{W}, \mathrm{L}, \mathrm{R}\}^4| = |\{\mathrm{W}, \mathrm{L}, \mathrm{R}\}|^4 = 3^4 = 3 \cdot 3 \cdot 3 \cdot 3 = (3 \cdot 3) \cdot (3 \cdot 3) = 9 \cdot 9 = 81$$

3-14. 200 people attended a party.

- 100 of them were at least 20 years old or older
- 150 of them were at most 30 years old
- How many of them were between 20 and 30 years old?

4. Random Variables

4-1. A fair coin is tossed twice. Let X denote the number of heads occurring.

(a) What is the range of X?
(b) What is the probability distribution of X? Give a table.
(c) Find $E[X]$.
(d) Find $\mathrm{Var}[X]$.
(e) Find σ_X (the standard deviation of X)

4-2. A fair coin is tossed until a heads or three tails occur. Let X denote the number of coin tosses.

(a) What is the range of X?
(b) What is the probability distribution of X? Give a table.
(c) Find $E[X]$.
(d) Find $\mathrm{Var}[X]$.

(e) Find σ_X (the standard deviation of X)

4-3. A player tosses two fair coins. The player wins \$3 if 2 heads occur, wins \$1 if 1 heads occurs, and **loses** \$5 if no heads occur. Let X denote the amount the player receives from playing the game once.

(a) What is the range of X?

(b) What is the probability distribution of X? Give a table.

(c) Find $E[X]$.

(d) Find $\text{Var}[X]$.

(e) Find σ_X (the standard deviation of X)

4-4. A single die is rolled. Let X be a random variable denoting the number facing up on the die, and let $Z = 2X$.

(a) What is the range of X?

(b) What is the probability distribution of X? Give a table.

(c) Find $E[X]$.

(d) Find $\text{Var}[X]$.

(e) What is the range of Z?

(f) What is the probability distribution of Z? Give a table.

(g) Find $E[Z]$.

(h) Find $\text{Var}[Z]$.

4-5. A single die is rolled. Let $Y = 1$ if the number facing up is odd, 0 if the number facing up is even.

(a) What is the range of Y?

(b) What is the probability distribution of Y? Give a table.

(c) Find $E[Y]$.

(d) Find $\text{Var}[Y]$.

5. Binomial and Normal Distributions

5-1. Look up the following values in Table A-1

(a) 0

(b) 0.01

(c) 0.5

(d) 0.57

(e) 1

(f) 1.41

(g) 1.414

(h) 1.5

(i) 2

(j) 3

(k) 3.14

(l) 3.99

(m) 4.0

(n) 5.0

5-2. What is the (approximate) area under the standard number curve between the each of the following pairs of values:

(a) 0 and 1

(b) 1 and 2

(c) 2 and 3

(d) 3 and 4
(e) 4 and 5
(f) -1 and 0
(g) -2 and 1
(h) -3 and -2
(i) -4 and -3
(j) -5 and -4
(k) -1 and 1
(l) -2 and 2
(m) -2 and 3
(n) -3 and 1
(o) -2 and 4

5-3. 420 6-sided dice are rolled. Let X denote the sum of the numbers facing up on the dice.

(a) Find $E[X]$
(b) Find $\mathrm{Var}[X]$
(c) Find σ_X
(d) What is $\Pr[1400 < X < 1500]$?
(e) What is $\Pr[1500 \leq X \leq 1600]$?
(f) What is $\Pr[1300 < X < 1400]$?
(g) What is $\Pr[X = 1500]$?

5-4. 180 6-sided dice are rolled. Let X denote the number of 5s rolled.

(a) Find $E[X]$
(b) Find $\mathrm{Var}[X]$
(c) Find σ_X
(d) What is $\Pr[25 < X < 30]$?
(e) What is $\Pr[25 < X \leq 30]$?
(f) What is $\Pr[X > 30]$?
(g) What is $\Pr[X \geq 30]$?
(h) What is $\Pr[X = 25]$?

5-5. A 6-sided die is rolled 400 times. Each time, Otto bets that the number facing up will be odd. Let X denote the number of times that Otto wins his bet.

(a) Find $E[X]$
(b) Find $\mathrm{Var}[X]$
(c) Find σ_X
(d) What is $\Pr[100 < X < 200]$?
(e) What is $\Pr[X < 200]$?
(f) What is $\Pr[200 \leq X \leq 300]$?
(g) What is $\Pr[X \geq 200]$?
(h) What is $\Pr[100 < X \leq 300]$?
(i) What is $\Pr[X = 200]$?
(j) What is $\Pr[X = 0]$? Give the exact answer, not just an estimate.
(k) What is $\Pr[X \geq 1]$? Give the exact answer, not just an estimate.
(l) What is $\Pr[X = 100]$?
(m) What is $\Pr[X = 200]$?
(n) What is $\Pr[X < 400]$? Give the exact answer, not just an estimate.
(o) What is $\Pr[X = 400]$? Give the exact answer, not just an estimate.

(p) What is $\Pr[X < 500]$? Give the exact answer, not just an estimate.

(q) What is $\Pr[X = 500]$? Give the exact answer, not just an estimate.

5-6. For this problem we will use a 3-sided die. When you roll a 3-sided die, the outcome is a 1, 2, or 3, and each outcome is equally likely. You can find pictures at
`http://www.advancinghordes.com/index.php/cPath/2_61_82`

Let X_1 denote the result of a single 3-sided die roll, and let X be the sum of 600 3-sided die rolls.

(a) Find $E[X_1]$

(b) Find $\text{Var}[X_1]$

(c) Find σ_{X_1}

(d) What is $\Pr[X_1 = 2]$?

(e) Find $E[X]$

(f) Find $\text{Var}[X]$

(g) Find σ_X

(h) What is $\Pr[1185 \leq X < 1225]$?

(i) What is $\Pr[1215 < X < 1225]$?

(j) What is $\Pr[1185 \leq X \leq 1215]$?

(k) What is $\Pr[X = 1200]$?

5-7. For this problem we will use a 4-sided die. In case you have never seen one, Wikipedia has a picture. You can see it by typing the following url into your browser:
`http://upload.wikimedia.org/wikipedia/commons/1/19/4-sided_dice_250.jpg` or you can search google images for "4-sided die." When you roll a 4-sided die, the outcome is the number on the bottom of the die, since there is no room to write a number on the top.

Let X_1 denote the result of a single 4-sided die roll, and let X be the sum of 80 4-sided die rolls.

(a) Find $E[X_1]$

(b) Find $\text{Var}[X_1]$

(c) Find σ_{X_1}

(d) What is $\Pr[X_1 = 2.5]$?

(e) Find $E[X]$

(f) Find $\text{Var}[X]$

(g) Find σ_X

(h) What is $\Pr[185 < X \leq 225]$?

(i) What is $\Pr[215 \leq X \leq 225]$?

(j) What is $\Pr[185 < X < 215]$?

(k) What is $\Pr[X = 200]$?

5-8. The ranking of players at a certain bridge club event has mean 17 and variance 4. (To be precise, let Z denote the ranking of a single randomly chosen player at the event. Then $E[Z] = 17$ and $\text{Var}[Z] = 4$.)

Bill selects 10 of the players at random, adds up their rankings, and calls the result X.

Carolyn selects 5 teams (consisting of 2 players each), adds up their rankings, and calls the result Y. You should assume that players tend to team up with players who have similar ranks.

(a) Find $E[X]$

(b) Find $\text{Var}[X]$

(c) Find $E[Y]$

(d) Is $\text{Var}[Y]$ equal to $\text{Var}[X]$, less than $\text{Var}[X]$, or greater than $\text{Var}[X]$?

5-9. A fair coin is tossed 1600 times. Let X denote the number of heads occurring.

(a) What is the range of X?

(b) What is the probability distribution of X **called**? Do not give a table.

(c) Find $E[X]$.
(d) Find $\mathrm{Var}[X]$.
(e) Find σ_X (the standard deviation of X)
(f) *Estimate $\Pr(X > 830)$
(g) *Estimate $\Pr(X \geq 830)$
(h) *Estimate $\Pr(X < 830)$
(i) *Estimate $\Pr(X \leq 830)$

5-10. A certain biased coin comes up heads with probability $\frac{1}{3}$ when spun. The coin is spun 18 times. Let X denote the number of heads occurring.
(a) What is the range of X?
(b) What is the probability distribution of X **called**? Do not give a table.
(c) Find $E[X]$.
(d) Find $\mathrm{Var}[X]$.
(e) Find σ_X (the standard deviation of X)
(f) Find $\Pr(X = 0)$.
(g) Find $\Pr(X \geq 1)$.
(h) Find $\Pr(X = 18)$.
(i) *Estimate $\Pr(X > 5)$
(j) *Estimate $\Pr(X \geq 5)$
(k) *Estimate $\Pr(X < 5)$
(l) *Estimate $\Pr(X \leq 5)$

5-11. A roulette wheel has 2 green pockets, 18 red pockets, and 18 black pockets which a ball can land. Each pocket is equally likely. A roulette wheel is spun 722 times. Let X denote the number of times the ball lands in a green pocket.
(a) What is the range of X?
(b) What is the probability distribution of X **called**? Do not give a table.
(c) Find $E[X]$.
(d) Find $\mathrm{Var}[X]$.
(e) Find σ_X (the standard deviation of X)
(f) Find $\Pr(X = 0)$.
(g) Find $\Pr(X \geq 1)$.
(h) *Estimate $\Pr(X > 40)$
(i) *Estimate $\Pr(X \geq 40)$
(j) *Estimate $\Pr(X < 40)$
(k) *Estimate $\Pr(X \leq 40)$

6. Sampling: Sample Mean and Sample Proportion

6-1. Joe's grandfather has 105,337 baseball cards in his collection. The average value of a card in the collection is \$5.75. The standard deviation is \$0.50. Joe picks 25 cards at random from the collection and computes their average value V.
(a) What is $E[V]$?
(b) What is σ_V?
(c) Estimate $\Pr[V \leq \$5.50]$
(d) Estimate $\Pr[\$5.50 \leq V \leq \$6.00]$
(e) Estimate $\Pr[V \leq \$6.00]$
(f) Estimate $\Pr[V = \$5.75]$

Note: V is not integer-valued, so do not add or subtract 0.5 from the endpoints of your interval.

6-2. 80% of all Vulcans oppose the war in Iraq. Kirk asks a random Vulcan whether he/she opposes the war in Iraq. Let $K = 1$ if Kirk's Vulcan says yes, 0 if that Vulcan says no. Spock asks 64 Vulcans whether they oppose the war in Iraq. Let S be the percentage of Vulcans in Spock's sample who say yes.

(a) What is $E[K]$?

(b) What is σ_K?

(c) What is $E[S]$?

(d) What is σ_S?

(e) Estimate $\Pr[S \leq 70\%]$

(f) Estimate $\Pr[70\% \leq S \leq 80\%]$

(g) Estimate $\Pr[S \leq 80\%]$

Note: S is not integer-valued, so do not add or subtract 0.5 from the endpoints of your interval.

6-3. A 3-sided die is rolled 600 times and the results are averaged. Call this average X.

(a) Find $E[X]$

(b) Find σ_X

(c) What is $\Pr[1.975 \leq X \leq 2.025]$?

(d) What is $\Pr[X = 2]$?

7. Sampling: Confidence Intervals

7-1. The College Board designed a new test and gave it to 2500 random students. Their scores averaged 550, with standard deviation 40.

(a) Give a 99% confidence interval for the mean test score in the entire population.

(b) How large a sample would you need in order to get a margin of error of 2.0 with 99% confidence?

7-2. Your company has been using an opinion poll with sample size 5000. The poll's margin of error is 6%. Your boss wants to improve the margin of error to 2%. What sample size should she use?

7-3. You have been assigned to predict the results of a 2-candidate election. What sample size would you need in order to ensure a 2.5% margin of error with 90% confidence?

7-4. The College Board designed a new test and gave it to 1600 random students. Their scores averaged 700, with standard deviation 60.

(a) What is the confidence level for the interval $[697, 703]$?

(b) Give a 98% confidence interval for the mean test score in the entire population.