innovate achieve lead

BITS Pilani

Pilani Campus

MATH F113
Probability and Statistics

Dr. Shivi Agarwal
Department of Mathematics

Chapter # 5

Joint Distributions

Single Random Variables: (Univariate)

Discrete

Continuous

Two Dimensional Random variables (Bivariate)

Continuous

Discrete

Bivariate distribution occurs when we observe 2 nondeterministic quantities, one followed by another.

For example :

1) Record the atmospheric temperature T in celsius followed by atmospheric pressure P in pounds per square foot at a random place and time. This gives 2 dimensional r.v. (T, P).

2) For a randomly chosen student, record number of A grades he has received followed by the number of his B grades in the last semester.

n-dimensional random variable $\leftrightarrow$ we observe n nondeterministic quantities in sequence.

Joint Probability Mass function

Let X and Y be discrete r.v defined on the sample space of an experiment, the ordered pair (X,Y) is called a two dimensional discrete r.v, A function

$$p_{XY}(x,y) = P[X = x \; and \; Y = y]$$

is called the joint probability mass function for each pair of (X,Y).

BITS Pilani, Pilani Campus

Necessary and Sufficient Conditions

$$1.\ p_{XY}(x,y) \geq 0 \quad \forall (x,y) \in R^2$$

$$2.\ \sum_{all\ x} \sum_{all\ y} p_{XY}(x,y) = 1$$

Two Discrete Random Variables

Now let A be any set consisting of pairs of (x, y) values (e.g., $A = \{(x, y): x + y = 5\}$ or $\{(x, y): \max(x, y) \leq 3\}$).

Then the probability $P[(X, Y) \in A]$ is obtained by summing the joint pmf over pairs in A:

$$P[(X, Y) \in A] = \sum_{(x,\, y)} \sum_{\in A} p_{XY}(x, y)$$

Example: In an automobile plant two tasks are performed by robots. The first entails welding 2 joints; the second, tightening 3 bolts. Let **X** denote the number of defective welds and **Y** the number of improperly tightened bolts produced per car. Since X and Y are each discrete, (X,Y) is a 2-dimensional discrete random variable.

Past data indicates that the joint probability mass function for (X,Y) is shown in Table. Note that each entry in the table is a number between 0 and 1 and therefore can be interpreted as a probability.

x \ y	0	1	2	3
0	0.840	0.030	0.020	0.010
1	0.060	0.010	0.008	0.002
2	0.010	0.005	0.004	0.001

Note : Sum of all entries =1

Marginal pmf

Let (X,Y) be a two dimensional discrete random variable with joint pmf p_{XY}

The marginal probability mass function of *X*

$$p_X(x) = \sum_{all\ y:\, p_{XY}(x,y)>0} p_{XY}(x,y)$$

The marginal probability mass function of Y

$$p_Y(y) = \sum_{all\ x:\, p_{XY}(x,y)>0} p_{XY}(x,y)$$

Marginal pmf

x \ y	0	1	2	3
0	0.840	0.030	0.020	0.010
1	0.060	0.010	0.008	0.002
2	0.010	0.005	0.004	0.001

Marginal Distributions

- Probability mass function of X alone: Sum the joint probability mass function over all values of Y

- Probability mass function of Y alone: Sum the joint probability mass function over all values of X

Q.1.

A service station has both self-service and full-service islands. On each island, there is a single regular unleaded pump with two hoses. Let X denote the number of hoses being used on the self-service island at a particular time, and let Y denote the number of hoses on the full-service island in use at that time. The joint pmf of X and Y appears in the accompanying tabulation.

Q.1.

$p_{XY}(x,y)$		y		
		0	1	2
x	0	0.10	0.04	0.02
	1	0.08	0.20	0.06
	2	0.06	0.14	0.30

a. What is $P(X = 1 \text{ and } Y = 1)$?

b. Compute $P(X \leq 1 \text{ and } Y \leq 1)$.

c. Give a word description of the event $\{X \neq 0 \text{ and } Y \neq 0\}$, and compute the probability of this event.

d. Compute the marginal pmf of X and of Y. Using $p_X(x)$, what is $P(X \leq 1)$?

Two Continuous Random Variables

Two Continuous Random Variables

The probability that the observed value of a continuous rv X lies in a one-dimensional set A (such as an interval) is obtained by integrating the pdf $f(x)$ over the set A.

Two Continuous Random Variables

Similarly, the probability that the pair (X, Y) of continuous rv's falls in a two-dimensional set A (such as a rectangle) is obtained by integrating a function called the *joint density function.*

Two Continuous Random Variables

Let X and Y be continuous rv's. **A joint probability density function** $f_{XY}(x, y)$ for these two variables is a function satisfying

$$1.\ f_{XY}(x,y) \geq 0 \ \forall (x,y) \in R^2$$

$$2. \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x,y)\ dy\ dx = 1$$

Two Continuous Random Variables

Then for any two-dimensional set A

$$P[(X, Y) \in A] = \iint_A f_{XY}(x, y)\ dy\ dx$$

In particular, if A is the two-dimensional rectangle $\{(x, y): a \le x \le b, c \le y \le d\}$, then

$$P[a \le X \le b \ and \ c \le Y \le d]$$

$$= \int_a^b \int_c^d f_{XY}(x, y)\ dy\ dx,$$ **for $a \le b, c \le d$ real**

Note: In one dimensional continuous case, the probabilities correspond to areas under density curve while in the case of 2-D, they corresponds to volumes under density surfaces.

Two Continuous Random Variables

Let (X,Y) be a two dimensional continuous random variable with joint density f_{XY}.

The marginal density for X, denoted by f_X is

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dy$$

Similarly define f_Y, the marginal density for Y.

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dx$$

Verify f_X, f_Y are densities.

Characterizing properties of densities?

Q. 9

Each front tire on a particular type of vehicle is supposed to be filled to a pressure of 26 psi. Suppose the actual air pressure in each tire is a random variable—*X* for the right tire and *Y* for the left tire, with joint pdf

$$f_{XY}(x,y)=\begin{cases} k(x^2+y^2) & 20 \le x \le 30, 20 \le y \le 30, \\ 0 & otherwise \end{cases}$$

a. What is the value of *k*?

b. What is the probability that both tires are underfilled?

Q.9.

$$f_{XY}(x,y)=\begin{cases} k(x^2+y^2) & 20 \le x \le 30, 20 \le y \le 30, \\ 0 & otherwise \end{cases}$$

d. Determine the (marginal) distribution of air pressure in the right tire alone.

e. Determine the (marginal) distribution of air pressure in the left tire alone.

c. What is the probability that the difference in air pressure between the two tires is at most 2 psi?

Independent Random Variables

Def: Two random variables X and Y are said to be independent if and only if

$$\textit{for all } x \textit{ and } y,$$
$$p(x, y) = p_X(x)\, p_Y(y) \quad \text{when X and Y are discrete}$$
$$f_{XY}(x, y) = f_X(x)\, f_Y(y) \quad \text{when X and Y are continuous}$$

If above condition is not satisfied for all (x, y), then X and Y are said to be **dependent.**

Independent Random Variables

Independence of two random variables is most useful when the description of the experiment under study suggests that X and Y have no effect on one another.

Then once the marginal pmf's or pdf's have been specified, the joint pmf or pdf is simply the product of the two marginal functions. It follows that

$$P\,(a \le X \le b,\, c \le Y \le d) = P\,(a \le X \le b) \cdot P\,(c \le Y \le d)$$

Q.1.

$p_{XY}(x,y)$		y			
		0	1	2	$p_X(x)$
x	0	0.10	0.04	0.02	0.16
	1	0.08	0.20	0.06	0.34
	2	0.06	0.14	0.30	0.50
	$p_Y(y)$	0.24	0.38	0.38	1.00

e. Are X and Y independent RV's? Explain

Q.2.

When an automobile is stopped by a roving safety patrol, each tire is checked for tire wear, and each headlight is checked to see whether it is properly aimed. Let X denote the number of headlights that need adjustment, and let Y denote the number of defective tires.

a. If X and Y are independent with $p_X(0) = .5$, $p_X(1) = .3$, $p_X(2) = .2$, and $p_Y(0) = .6$, $p_Y(1) = .1$, $p_Y(2) = p_Y(3) = .05$, and $p_Y(4) = .2$,

Q.2.

a. Display the joint pmf of (X, Y) in a joint probability table.

b. Compute $P(X \leq 1 \text{ and } Y \leq 1)$ from the joint probability table, and verify that it equals the product $P(X \leq 1)$. $P(Y \leq 1)$.

c. What is $P(X + Y = 0)$ (the probability of no violations)?

d. Compute $P(X + Y \leq 1)$.

Q.17.

An ecologist wishes to select a point inside a circular sampling region according to a uniform distribution (in practice this could be done by first selecting a direction and then a distance from the center in that direction). Let X the x coordinate of the point selected and Y the y coordinate of the point selected. If the circle is centered at (0, 0) and has radius R, then the joint pdf of X and Y is

Q.17.

$$f_{XY}(x,y)=\begin{cases}\dfrac{1}{\pi R^2}; & x^2+y^2\le R^2\\ 0\quad ; & otherwise\end{cases}$$

a. What is the probability that the selected point is within *R*/2 of the center of the circular region? [*Hint*: Draw a picture of the region of positive density *D*. Because $f_{XY}(x, y)$ is constant on *D*, computing a probability reduces to computing an area.]

Q.17.

b. What is the probability that both X and Y differ from 0 by at most $R/2$?

c. Answer part (b) for $R/\sqrt{2}$ replacing $R/2$.

d. What is the marginal pdf of X ? of Y? Are X and Y independent?

n-dimensional random variables

- The concept can be extended in an analogous manner to n dimensions.
- Replicate the definition.

More than two Random Variables

If $X_1, X_2, \ldots, X_n$ are all discrete random variables, the joint pmf of the variables is the function

$$p(x_1, x_2, \ldots, x_n) = P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$

More than two Random Variables

If the variables are continuous, the joint pdf of $X_1, \ldots, X_n$ is the function $f(x_1, x_2, \ldots, x_n)$ such that for any n intervals $[a_1, b_1], \ldots, [a_n, b_n]$,

$$P(a_1 \le X_1 \le b_1, \ldots, a_n \le X_n \le b_n) = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, \ldots, x_n)\, dx_n \ldots dx_1$$

Multinomial distribution

In a binomial experiment, each trial could result in one of only two possible outcomes.

Consider now an experiment consisting of n independent and identical trials, in which each trial can result in any one of r possible outcomes.

Multinomial distribution

Let $p_i = P$ (outcome i on any particular trial), and define random variables by X_i = the number of trials resulting in outcome i ($i = 1, \ldots, r$).

Such an experiment is called a **multinomial experiment,** and the joint pmf of $X_1, \ldots, X_r$ is called the **multinomial distribution.**

Multinomial distribution

By using a counting argument analogous to the one used in deriving the binomial distribution, the joint pmf of $X_1, \ldots, X_r$ can be shown to be

$$p(x_1, \ldots, x_r)$$

$$= \begin{cases} \dfrac{n!}{(x_1!)(x_2!) \cdot \cdots \cdot (x_r!)} p_1^{x_1} \cdot \ldots \cdot p_r^{x_r} & x_i = 0, 1, 2, \ldots, \text{with } x_1 + \cdots + x_r = n \\ 0 & \text{otherwise} \end{cases}$$

Multinomial distribution

$$p(x_1, \ldots, x_r) = \begin{cases} \dfrac{n!}{(x_1!)(x_2!) \cdot \cdots \cdot (x_r!)} p_1^{x_1} \cdot \ldots \cdot p_r^{x_r} & x_i = 0, 1, 2, \ldots, \text{ with } x_1 + \cdots + x_r = n \\ 0 & \text{otherwise} \end{cases}$$

The case $r = 2$ gives the binomial distribution, with

X_1 = number of successes and
$X_2 = n - X_1$ = number of failures.

innovate achieve lead

Q. 20.

Let X_1, X_2, X_3, X_4, X_5, and X_6 denote the numbers of blue, brown, green, orange, red, and yellow M&M candies, respectively, in a sample of size n. Then these X_i's have a multinomial distribution. According to the M&M Web site, the color proportions are $p_1 = .24$, $p_2 = .13$, $p_3 = .16$, $p_4 = .20$, $p_5 = .13$, and $p_6 = .14$.

Q. 20.

a. If n=12, what is the probability that there are exactly two M&Ms of each color?

b. For n=20, what is the probability that there are at most five orange candies?

c. In a sample of 20 M&Ms, what is the probability that the number of candies that are blue, green, or orange is at least 10?

Independent Random Variables

In general for n random variables, $X_1, \ldots, X_n$ are indep iff

$$p_{X_1 \ldots X_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} p_{X_i}(x_i)$$

$$f_{X_1 \ldots X_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i)$$

the joint pmf or pdf of the subset is equal to the product of the marginal pmf's or pdf's.

Cumulative distribution function

Def: Let (X,Y) be a two dimensional random variable. The cdf **F** of the two dimensional random variable (X,Y) is defined by:

$$F(x,y) = P[X \le x \text{ and } Y \le y]$$

If F is the cdf of a two dimensional random variable with joint density, f, then

$$\frac{\partial^2}{\partial x\, \partial y} F(x, y) = f(x, y), \text{wherever}$$

$$F \text{ is differentiable in case}$$

$$\text{of continuous case.}$$

Conditional Distribution

The conditional pmf for X given Y= y for discrete joint random variable (X,Y), denoted by $p_{X|Y}$,

$$p_{X|Y}(x \mid y) = \frac{P[X = x \text{ and } Y = y]}{P[Y = y]}$$

$$= \frac{p_{XY}(x, y)}{p_Y(y)}, \text{ provided } p_Y(y) \neq 0.$$

Conditional Distribution

The conditional pmf for Y given X= x for discrete joint random variable (X,Y), denoted by $p_{Y|X}$,

$$p_{Y|X}(y \mid x) = \frac{P[X = x \ and \ \ Y = y]}{P[X = x]}$$

$$= \frac{p_{XY}(x, y)}{p_X(x)}, \text{ provided } p_X(x) \neq 0.$$

Def: Let (X,Y) be a two dimensional continuous random variable with joint pdf f_{XY} and marginal pdfs f_X and f_Y. Then

1. The conditional density for X given Y= y, denoted by $f_{X|Y}$ is given by:

$$f_{X|Y}(x \mid y) = \frac{f_{XY}(x, y)}{f_Y(y)}, \; f_Y(y) > 0$$

2. The conditional density for Y given X= x, denoted by $f_{Y|X}$ is given by:

$$f_{Y|X}(y \mid x) = \frac{f_{XY}(x, y)}{f_X(x)}, \; f_X(x) > 0$$

Note: The conditional densities satisfies all the requirements for a one dimensional pdf. Thus for a fixed y with $f_Y(y) \neq 0$, we have $f_{X|Y}(x|y) \geq 0$ for all x and

$$\int_{-\infty}^{\infty} f_{X|Y}(x \mid y)\, dx = \int_{x=-\infty}^{\infty} \frac{f(x, y)}{f_Y(y)}\, dx$$

$$= \frac{1}{f_Y(y)} \int_{x=-\infty}^{\infty} f(x, y)\, dx = \frac{f_Y(y)}{f_Y(y)} = 1$$

Q. 18.

$p_{XY}(x,y)$		y			
		0	1	2	$p_X(x)$
x	0	0.10	0.04	0.02	0.16
	1	0.08	0.20	0.06	0.34
	2	0.06	0.14	0.30	0.50
	$p_Y(y)$	0.24	0.38	0.38	1.00

Refer to Exercise 1 and answer the following questions:

a. Given that $X=1$, determine the conditional pmf of Y—i.e., $p_{Y|X}(0|1)$, $p_{Y|X}(1|1)$, and $p_{Y|X}(2|1)$.

Q. 18.

b. Given that two hoses are in use at the self-service island, what is the conditional pmf of the number of hoses in use on the full-service island?

c. Use the result of part (b) to calculate the conditional probability $P(Y \leq 1 \mid X=2)$.

d. Given that two hoses are in use at the full-service island, what is the conditional pmf of the number in use at the self-service island?

Q. 21.

The joint pdf of pressures for right and left front tires is given in Exercise 9.

$$f_{XY}(x,y)=\begin{cases} k(x^2+y^2) & 20 \le x \le 30, 20 \le y \le 30, \\ 0 & otherwise \end{cases}$$

a. Determine the conditional pdf of Y given that $X = x$ and the conditional pdf of X given that $Y = y$.

Q. 21.

b. If the pressure in the right tire is found to be 22 psi, what is the probability that the left tire has a pressure of at least 25 psi? Compare this to $P(Y \geq 25)$.

c. If the pressure in the right tire is found to be 22 psi, what is the expected pressure in the left tire, and what is the standard deviation of pressure in this tire?

5.2. Expectation, Covariance, and Correlation

- Any function $h(X)$ of a single rv X is itself a random variable.
- $E[h(X)]$ is computed as a weighted average of $h(x)$ values, where the weight function is the pmf $p(x)$ or pdf $f(x)$ of X.
- A similar result holds for a function $h(X, Y)$ of 2-D jointly distributed random variables.

Expectation

Def: Let (X,Y) be a 2-D discrete r.v with joint pmf $p_{XY}(x,y)$. Let h(X,Y) be a real valued function of (X, Y).

The expected value of h(X,Y), denoted by E[h(X,Y)] is given by:

$$E[h(X,Y)] = \sum_{all\ x} \sum_{all\ y} h(x,y) . p_{XY}(x,y) = \mu_{h(X,Y)}$$

provided

$$\sum_{all\ x} \sum_{all\ y} | h(x,y) | . p_{XY}(x,y) \quad \text{converges.}$$

Expectation

For (X,Y) continuous rv with pdf $f_{XY}(x,y)$

$$E[h(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y).\, f_{XY}(x,y)dxdy$$

provided

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\, h(x,y)\,|\,.\, f_{XY}(x,y)\ dy\, dx \text{ exists.}$$

Univariate Average found via the joint distribution

$$E[X] = \sum_{all\ x}\sum_{all\ y} x.\ p_{XY}(x,y) \quad for\,(X,Y)\ discrete$$

$$E[Y] = \sum_{all\ x}\sum_{all\ y} y.\ p_{XY}(x,y) \quad for\,(X,Y)\ discrete$$

$$E[X] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x.f_{XY}(x,y)\ dy\,dx$$

$$E[Y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y.f_{XY}(x,y)\ dx\,dy$$

$$for\ \ (X,Y)\ continuous$$

Variance

- Var[h(X,Y)]=E[(h(X,Y)-E[h(X,Y)])^2] if all expectations considered exist.
- Recall : E[aX+bY+c]=aE[X]+bE[Y]+c.
- More generally, for n-dimensional r.v.

$$E[a_0 + a_1 X_1 + \ldots + a_n X_n]$$
$$= a_0 + a_1 E[X_1] + \ldots + a_n E[X_n]$$

Q. 26.

Consider a small ferry that can accommodate cars and buses. The toll for cars is \$3, and the toll for buses is \$10. Let X and Y denote the number of cars and buses, respectively, carried on a single trip. Suppose the joint distribution of X and Y is as given in the table on next slide. Compute the expected revenue from a single trip.

Q. 26.

$p_{XY}(x,y)$		y		
		0	1	2
x	0	0.025	0.015	0.010
	1	0.050	0.030	0.020
	2	0.125	0.075	0.050
	3	0.150	0.090	0.060
	4	0.100	0.060	0.040
	5	0.050	0.030	0.020

Q. 31.

Annie and Alvie have agreed to meet for lunch between noon (0:00 P.M.) and 1:00 P.M. Denote Annie's arrival time by X, Alvie's by Y, and suppose X and Y are independent with pdf's given on next slide.

What is the expected amount of time that the one who arrives first must wait for the other person?

Q. 31.

$$f_X(x) = \begin{cases} 3x^2 & ; 0 \le x \le 1 \\ 0 & ; \ otherwise \end{cases}$$

$$f_Y(y) = \begin{cases} 2y & ; 0 \le y \le 1 \\ 0 & ; \ otherwise \end{cases}$$

Covariance

When two random variables X and Y are not independent, it is frequently of interest to assess how strongly (linear) they are related to one another.

Covariance

Def: Let X and Y be random variables with means μ_X and μ_Y respectively. The Covariance between X and Y, denoted by Cov(X,Y) or σ_{XY} is given by:

$$\mathrm{Cov}(X, Y) = E[(X - \mu_x)(Y - \mu_Y)]$$

Remark: If small values of X tend to be associated with small values of Y and large values of X with large values of Y, Then $(X-\mu_X)(Y-\mu_Y) > 0$, yields a positive covariance and the reverse is true, that is small values of X tend to associated with large values of Y and vice versa, then $(X-\mu_X)(Y-\mu_Y) < 0$ yields a negative covariance. In essence covariance is an indication of how X and Y are related.

If X and Y are not linearly related, positive and negative products will tend to cancel one another, yielding a covariance near 0.

Theorem: $\text{Cov}(X,Y) = E[XY] - E[X]\ E[Y]$

Q. 28.

Let (X,Y) be a two dimensional random variable. If X and Y are independent then $E[XY] = E[X]\,E[Y]$.

If X, Y are independent, then Cov(X,Y)=0.

But the **Converse is not true**: that is, 'zero covariance implies independence' is **not true**.

EX : $f_{XY}(x,y) = 1$ if $|x|<1$ and $|y|<|x|/2$

$f_{XY}(x,y)=0$ otherwise.

Show : $Cov(X,Y) = 0$ but X, Y are not indep.

Note: If $X_i, i = 1, \ldots, n$ are random va

and $a_i, i = 0, 1 \ldots, n$ are constants then

$$Var\left[a_0 + \sum_{i=1}^{n} a_i X_i\right]$$

$$= \sum_{i=1}^{n} a_i^2 Var[X_i] + 2\sum_{i<j} a_i a_j Cov[X_i, X_j]$$

In particular, if $X_1, \ldots, X_n$ are independent, then

$$Var\left[a_0 + \sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2 Var[X_i]$$

Prove that

$$Var(X+Y) = Var(X) + Var(Y) + 2\, Cov(X,Y)$$

Show that if $X=Y$, then

$$Cov(X,Y) = Var(X) = Var(Y)$$

Correlation

The **correlation coefficient** of X and Y, denoted by Corr(X, Y), $\rho_{X,Y}$, or just ρ, is defined by

$$\rho_{X,Y} = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \cdot \sigma_Y}$$

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 $a \neq 0$.

Correlation

- ρ is a measure of the degree of **linear** relationship between X and Y,

- ρ <1 in absolute value indicates only that the relationship is not completely linear, but there may still be a very strong nonlinear relation.

Correlation

- Also, $\rho = 0$ does not imply that X and Y are independent, but only that there is a complete absence of a linear relationship. When $\rho = 0$, X and Y are said to be **uncorrelated.**
- Two variables could be uncorrelated yet highly dependent because there is a strong nonlinear relationship, so be careful not to conclude too much from knowing that $\rho = 0$.

Q. 23.

a. Compute the covariance between *X* and *Y* in Exercise 9.

b. Compute the correlation coefficient ρ for this *X* and *Y*.

$$f_{XY}(x,y)=\begin{cases} k(x^2+y^2) & 20\le x\le 30, 20\le y\le 30, \\ 0 & otherwise \end{cases}$$

Q. 35.

a. Use the rules of expected value to show that $\text{Cov}(aX+b, cY+d) = ac\,\text{Cov}(X, Y)$.

b. Use part (a) along with the rules of variance and standard deviation to show that $\text{Corr}(aX+b, cY+d) = \text{Corr}(X, Y)$ when a and c have the same sign.

c. What happens if a and c have opposite signs?

Q. 36.

Show that if $Y = aX + b$ ($a \neq 0$), then Corr(X, Y) = +1 or -1. Under what conditions will $\rho = +1$?

Q. 87.

a. Use the general formula for the variance of a linear combination to write an expression for $V(aX+Y)$. Then let $a=\sigma_Y/\sigma_X$, and show that $\rho \geq -1$.

b. By considering $V(aX-Y)$, conclude that $\rho \leq 1$.

c. Use the fact that $V(W) = 0$ only if W is a constant to show that $\rho = 1$ only if $Y = aX + b$.