In this case, $X_B = 0$ and $X_A = Y_A$ has a normal distribution, and we can standardize the probability. $P[X_B > X_A|D_4] = P[Y_B > Y_A] = P[Y_B - Y_A > 0]$. Since claims from the two companies are independent, Y_A and Y_B are independent. The sum or difference of normal random variables is normal, and the mean is the sum or difference of the means. The mean of $Y_B - Y_A$ is $9,000 - 10,000 = -1,000$. Since Y_A and Y_B are independent,

$$Var[Y_B - Y_A] = 2000^2 + 2000^2 = 8,000,000$$

Then standardizing $Y_B - Y_A$, we get

$$P[X_B > X_A|D_4] = P[Y_B > Y_A] = P[Y_B - Y_A > 0]$$
$$= P\left[\frac{Y_B - Y_A - (-1,000)}{\sqrt{8,000,000}} > \frac{0-(-1,000)}{\sqrt{8,000,000}}\right] = P[Z > 0.3536] = 1 - \Phi(0.35) = 0.363$$

(from the normal distribution table, we get $\Phi(0.35) = 0.5 \times \Phi(0.3) + 0.5 \times \Phi(0.4) = 0.637$).
Finally, $P[X_B > X_A] = 0 \times 0.42 + 1 \times 0.18) + 0 \times 0.28 + 0.363 \times 0.12 = 0.223$.

This solution can be summarized using some "general intuitive reasoning" as follows. The only way that X_B can be greater than X_A is if Company B has some claims. If Company A has no claims (prob. 0.6) and Company B has some claims (prob. 0.3), then the probability that Company B's claim amount will exceed Company A's claim is 1. If Company A has some claims (prob. 0.4) and Company B has some claims (prob. 0.3), then the probability that Company B's claims exceed Company A's claims is 0.363 (as outlined above). The overall probability that Company B's claims exceed Company A's claims is
$0.6 \times 0.3 \times 1 + 0.4 \times 0.3 \times 0.363) = .223$. Answer: D

9. Let T_g be the time until a claim from the good driver. Then the pdf of T_g is $f_g(t) = \frac{1}{6}e^{-t/6}$ (exponential with a mean of 6). Let T_b be the time until a claim from the bad driver. Then the pdf of T_b is $f_b(t) = \frac{1}{3}e^{-t/3}$ (exponential with a mean of 3).

Let A be the event that the first claim from a good driver will be filed within 3 years.
$P[A] = P[T_g < 3] = \int_0^3 \frac{1}{6}e^{-t/6}dt = 1 - e^{-1/2}$.

Let B be the event that the first claim from a bad driver will be filed within 2 years.
$P[B] = P[T_b < 2] = \int_0^2 \frac{1}{3}e^{-t/3}dt = 1 - e^{-2/3}$.

The probability that the first claim from a good driver will be filed within 3 years and the first claim from a bad driver will be filed within 2 years is $P[A \cap B]$. Since T_g and T_b are independent, so are the events A and B. Therefore, $P[A \cap B] = P[A] \times P[B] = (1 - e^{-1/2})(1 - e^{-2/3}) = 1 - e^{-1/2} - e^{-2/3} + e^{-7/6}$.
Answer: C

10. X has a $N(0,1)$ distribution, so that the density function of the conditional distribution is $f(x \mid X>0) = \frac{f(x)}{P[X>0]} = \frac{f(x)}{1/2} = 2f(x)$. The conditional expectation is $\int_0^\infty x \times f(x \mid X>0)\,dx = \int_0^\infty 2x \times \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx = -\frac{2}{\sqrt{2\pi}} e^{-x^2/2}\Big|_{x=0}^{x=\infty} = \frac{2}{\sqrt{2\pi}} = \sqrt{\frac{2}{\pi}}$

Answer: D

11. $E(E_1) = E(E_2) = 0$, $Var(E_1) = (0.0056h)^2$, $Var(E_2) = (0.0044h)^2$.
$E_1 + E_2$ is normal with $E(E_1 + E_2) = 0$, and since E_1 and E_2 are independent,
$Var(E_1+E_2) = Var(E_1) + Var(E_2) = (0.0056h)^2 + (0.0044h)^2 = (0.00712h)^2$.

The average of E_1 and E_2 is $A = \frac{1}{2}(E_1+E_2)$, which is also normal with mean 0 and variance $Var(A) = \frac{1}{4}Var(E_1+E_2) = (0.00356h)^2$.

The average height is within $.005h$ of the height of the tower if the absolute error in the average is less than $.005h$. We wish to find $P(|A| < 0.005h) = P(-0.005h < A < 0.005h)$.

We standardize A to get

$$P(-0.005h < A < 0.005h) = P\left(\frac{-0.005h - E(A)}{\sqrt{Var(A)}} < \frac{A - E(A)}{\sqrt{Var(A)}} < \frac{0.005h - E(A)}{\sqrt{Var(A)}}\right)$$
$$= P\left(\frac{-0.005h-0}{0.00356h} < Z < \frac{0.005h-0}{0.00356h}\right) = P(-1.4 < Z < 1.4) \ ,$$

where Z has a standard no0mal distribution. From the standard normal table,
$P(Z < 1.4) = 0.9192$, so that $P(Z < -1.4) = P(Z > 1.4) = 0.0808$,
and therefore, $P(-1.4 < Z < 1.4) = 0.9192 - 0.0808 = 0.8384$. Answer: D

12. The expected payout on the warranty is $\int_0^3 100(1-\frac{t}{3}) \times \frac{1}{n}\,dt = \frac{300}{2n} = 10 \rightarrow n = 15$. Answer: D

13. The exponential distribution with a mean of 2 has density function
$f(x) = \frac{1}{2}e^{-x/2}$, for $x > 0$, and distribution function $P[X \le x] = F(x) = 1 - e^{-x/2}$, for $x > 0$.

The probability that a printer will fail in the first year is $P[X \le 1] = F(1) = 1 - e^{-1/2} = 0.39347$ so that the expected number of failures in the first year out of 100 printers is 39.347.

The probability that a printer will fail in the second year is
$P[1 < X \le 2] = F(2) - F(1) = e^{-1/2} - e^{-2/2} = 0.23865$ so that the expected number of failures in the first year out of 100 printers is 23.865. The expected amount the manufacturer will pay in refunds is $(200)(39.347) + (100)(23.865) = 10,256$. Answer: D

14. The density function for the time of failure T is $f(t) = 0.1 \times e^{-.1t}$ (exponential with mean 10).

The amount paid is $P(t) = \begin{cases} x & 0 < t \le 1 \\ 0.5x & 1 < t \le 3 \\ 0 & t > 3 \end{cases}$. The expected amount paid is

$E[P(T)] = \int_0^{\infty} P(t) \times f(t)\, dx = \int_0^1 0.1xe^{-.1t}\, dt + \int_1^3 0.05xe^{-.1t}\, dt$
$= x[1 - e^{-.1}] + 0.5x[e^{-.1} - e^{-.3}] = 0.17717x.$

In order for this to be 1000, we must have $0.17717x = 1000 \rightarrow x = 5644.$ Answer: D

15. The exponential time until failure random variable T has density function of the form $f(t) = \frac{1}{\theta} e^{-t/\theta}$, and had distribution function $F(t) = P[T \le t] = 1 - e^{-t/\theta}$ for $t > 0$.
The median of the distribution is the time point m that satisfies the relationship $F(m) = 0.5$; in other words, m is the time point for which there is a 50% probability of failure by time m. We are given that $m = 4$, and therefore $F(4) = 1 - e^{-4/\theta} = 0.5$, from which it follows that $e^{-4/\theta} = 0.5$. We are asked to find $P[T > 5] = 1 - F(5) = e^{-5/\theta}$. Using the relationship $e^{-5/\theta} = (e^{-4/\theta})^{1.25}$, we get $P[T > 5] = e^{-5/\theta} = (.5)^{1.25} = 0.420$. Notice that we could solve for λ from the equation $e^{-4/\theta} = 0.5$, but it is not necessary. Answer: D

16. The coefficient of variation of a random variable X is $\frac{\sqrt{Var[X]}}{E[X]}$.
If X denotes the claim amount for the current one year period, then $E[X] = \mu$, $Var[X] = \mu^2$.

The claim amount for the one year period following the current one year period is $1.1X$, with mean $E[1.1X] = (1.1)E[X] = 1.1\mu$, and variance $Var[1.1X] = (1.1)^2 Var[X] = (1.1)^2\mu^2$.
The coefficient of variation in the following period is
$\frac{\sqrt{Var[1.1X]}}{E[1.1X]} = \frac{\sqrt{(1.1)^2 Var[X]}}{(1.1)E[X]} = \frac{\sqrt{Var[X]}}{E[X]} = \frac{\sqrt{\mu^2}}{\mu} = 1.$ Answer: C

17. The cdf for distribution 1 is $F_1(x) = 1 - e^{-x/100}$.
The median m_1 must satisfy $.5 = F_1(m_1) = 1 - e^{-m_1/100} \rightarrow m_1 = 69.3.$
The cdf for distribution 2 is $F_2(x) = \int_0^x f_2(t)\, dt = \int_0^x \frac{2\theta^2}{(t+\theta)^3}\, dt = 1 - \frac{\theta^2}{(x+\theta)^2}$.
The mean of distribution 2 is $E[X_2] = \int_0^{\infty} [1 - F_2(x)]\, dx = \int_0^{\infty} \frac{\theta^2}{(x+\theta)^2}\, dx = \theta = 100.$
Therefore, the cdf of distribution 2 is $F_2(x) = 1 - \frac{100^2}{(x+100)^2}$, and the median m_2 satisfies
$.5 = F_2(m_2) = 1 - \frac{100^2}{(m_2+100)^2} \Rightarrow m_2 = 41.4$. Then $\frac{m_1}{m_2} = \frac{69.3}{41.4} = 1.67.$
Answer: D

18. T = time, in years, until next major hurricane, is exponentially distributed with mean μ. The density function of T is $f(t) = \frac{1}{\mu}e^{-t/\mu}$, and cumulative distribution function is
$F(t) = P[T \le t] = 1 - e^{-t/\mu}$. We are given $P[T \le 10] = 1.5P[T \le 5]$, so that
$1 - e^{-10/\mu} = (1.5)[1 - e^{-5/\mu}]$. This can be written as $e^{-10/\mu} - 1.5e^{-5/\mu} + 0.5 = 0$.
With $x = e^{-5/\mu}$, this becomes the quadratic equation $x^2 - 1.5x + 0.5 = 0$ with roots
$x = 1$, 0.5. Therefore $e^{-5/\mu}$ is either 1 or 0.5. It is not possible to have $e^{-5/\mu} = 1$,
since that would require $\mu = \infty$. Therefore, $e^{-5/\mu} = 0.5$, so that $\mu = \frac{-5}{ln\frac{1}{2}} = \frac{5}{ln\,2}$. Answer: C

19. We denote by N the number of people in the car and by X the total loss due to hospitalization. We wish to find $E[N|X<1]$. The distribution of N is $P(N=0) = 0.7 \times 0.7 = 0.49$,
$P(N=1) = 2 \times .7 \times .3 = .42$ and $P(N=2) = .3 \times .3 = .09$.
Also, $P(X<1|N=0) = 1$, $P(X<1|N=1) = 1$, and
$P(X<1|N=2) = \int_0^1 f_{X_1}(t) \times P(X_2 < 1-t)\,dt$, where X_1 and X_2 are the two hospitalization costs.
This is $\int_0^1 1 \times (1-t)\,dt = \frac{1}{2}$.
To find $E[N|X<1]$, we first find the conditional probabilities $P(N=i|X<1) = \frac{P(N=i \cap X<1)}{P(X<1)}$.
We find $P(X<1) = P(X<1|N=0) \times P(N=0) + P(X<1|N=1) \times P(N=1)$
$+ P(X<1|N=2) \times P(N=2) = 1 \times 0.49 + 1 \times 0.42 + 0.5 \times 0.09 = 0.955$.
Then, to find $E[N|X<1]$, we find the conditional probabilities
$P(N=i|X<1) = \frac{P(N=i \cap X<1)}{P(X<1)} = \frac{P(N=i \cap X<1)}{0.955}$.
From the previous calculations, we have
$P(N=0 \cap X<1) = P(X<1|N=0) \times P(N=0) = 0.49$
$P(N=1 \cap X<1) = P(X<1|N=1) \times P(N=1) = 0.42$
$P(N=2 \cap X<1) = P(X<1|N=2) \times P(N=2) = 0.045$
Then, $E[N|X<1] = \sum_{i=0}^{2} i \times P(N=i|X<1)$
$= \frac{0 \times 0.49 + 1 \times 0.42 + 2 \times 0.045}{0.955} = 0.534$. Answer: B

20. $E[X] = \alpha\theta$ and $Var[X] = \alpha\theta^2$ so that $\alpha\theta = 2\alpha\theta^2$ from which we get $\theta = \frac{1}{2}$.
$M_X(t) = \left(\frac{1}{1-t\theta}\right)^\alpha$ so that $M_X(1) = \left(\frac{1}{1-\frac{1}{2}}\right)^\alpha = 2^\alpha = 16$ from which we get $\alpha = 4$.
Then, $E[X] = \frac{4}{2} = 2$. Answer: D

SECTION 8 - JOINT, MARGINAL, AND CONDITIONAL DISTRIBUTIONS

Joint distribution of random variables X and Y

A random variable X is a numerical outcome that results from some random experiment, such as the number that turns up when tossing a die. It is possible that an experiment may result in two or more numerical outcomes. A simple example would be the numbers that turn up when tossing two dice. X could be the number that turns up on the first die and Y could be the number on the second die. Another example could be the following experiment. A coin is tossed and if the outcome is head then toss one die, and if the outcome is tails then toss two dice. We could set $X = 1$ for a head and $X = 2$ for a tail and $Y =$ total on the dice thrown. In both of the examples just described, we have a pair of random variables X and Y, that result from the experiment. X and Y might be unrelated or independent of one another (as in the example of the toss of two independent dice), or they might be related to each other (as in the coin-dice example).

We describe the probability distribution of two or more random variables together as a **joint distribution**. As in the case of a single discrete random variable, we still describe probabilities for each possible pair of outcomes for a pair of discrete random variables. In the case of a pair of random variables X and Y, there would be probabilities of the form $P[(X = x) \cap (Y = y)]$ for each pair (x, y) of possible outcomes. For a pair of continuous random variables X and Y, there would be a density function to describe density over a two dimensional region.

A joint distribution of two random variables has a probability function or probability density function $f(x, y)$ that is a function of two variables (sometimes denoted $f_{X,Y}(x, y)$). It is defined over a two-dimensional region. For joint distributions of continuous random variables X and Y, the region of probability (the probability space) is usually a rectangle or triangle in the x-y plane.

If X and Y are discrete random variables, then $f(x, y) = P[(X = x) \cap (Y = y)]$ is the joint probability function, and it must satisfy

(i) $0 \leq f(x, y) \leq 1$ and

(ii) $\sum_x \sum_y f(x, y) = 1$

If X and Y are continuous random variables, then $f(x, y)$ must satisfy

(i) $f(x, y) \geq 0$ and

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

In the two dice example described above, if the two dice are tossed independently of one another then

$f(x, y) = P[(X = x) \cap (Y = y)] = P[X = x] \times P[Y = y] = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$ for each pair with $x = 1, 2, 3, 4, 5, 6$

and $y = 1, 2, 3, 4, 5, 6$. The coin-die toss example above is more complicated because the number of dice tossed depends on whether the toss is head or tails. If the coin toss is a head then $X = 1$ and $Y = 1, 2, 3, 4, 5, 6$ so

$$f(1, y) = P[(X = 1) \cap (Y = y)] = \tfrac{1}{2} \times \tfrac{1}{6} = \tfrac{1}{12} \text{ for } y = 1, 2, 3, 4, 5, 6$$

If the coin toss is tail then $X = 2$ and $Y = 2, 3, ..., 12$ with

$$f(2, 2) = P[(X = 2) \cap (Y = 2)] = \tfrac{1}{2} \times \tfrac{1}{36} = \tfrac{1}{72}$$
$$f(2, 3) = P[(X = 2) \cap (Y = 3)] = \tfrac{1}{2} \times \tfrac{2}{36} = \tfrac{1}{36} \text{ , etc.}$$

It is possible to have a joint distribution in which one variable is discrete and one is continuous, or either has a mixed distribution. The joint distribution of two random variables can be extended to a joint distribution of any number of random variables.

If A is a subset of two-dimensional space, then $P[(X, Y) \in A]$ is the summation (discrete case) or double integral (continuous case) of $f(x, y)$ over the region A.

Example 8-1:

X and Y are discrete random variables which are jointly distributed with the probability function $f(x, y)$ defined in the following table:

		X		
		-1	0	1
	1	$\frac{1}{18}$	$\frac{1}{9}$	$\frac{1}{6}$
Y	0	$\frac{1}{9}$	0	$\frac{1}{6}$
	-1	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{9}$

From this table we see, for example, that $P[X = 0, Y = -1] = f(0, -1) = \frac{1}{9}$.

Find (i) $P[X + Y = 1]$, (ii) $P[X = 0]$ and (iii) $P[X < Y]$.

Solution:

(i) We identify the (x, y)-points for which $X + Y = 1$, and the probability is the sum of $f(x, y)$ over those points. The only x, y combinations that sum to 1 are the points $(0, 1)$ and $(1, 0)$.
Therefore, $P[X + Y = 1] = f(0, 1) + f(1, 0) = \frac{1}{9} + \frac{1}{6} = \frac{5}{18}$.

(ii) We identify the (x, y)-points for which $X = 0$. These are $(0, -1)$ and $(0, 1)$ (we omit $(0, 0)$ since there is no probability at that point). $P[X = 0] = f(0, -1) + f(0, 1) = \frac{1}{9} + \frac{1}{9} = \frac{2}{9}$

(iii) The (x, y)-points satisfying $X < Y$ are $(-1, 0)$, $(-1, 1)$ and $(0, 1)$.
Then $P[X < Y] = f(-1, 0) + f(-1, 1) + f(0, 1) = \frac{1}{9} + \frac{1}{18} + \frac{1}{9} = \frac{5}{18}$. □

Example 8-2:

Suppose that $f(x, y) = K(x^2 + y^2)$ is the density function for the joint distribution of the continuous random variables X and Y defined over the unit square bounded by the points $(0, 0)$, $(1, 0)$, $(1, 1)$ and $(0, 1)$, find K. Find $P[X + Y \geq 1]$.

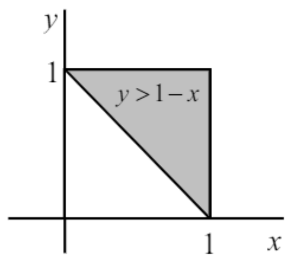
For trial use only.
Content may not
represent the most
current version

Solution:

In order for $f(x,y)$ to be a properly defined joint density, the (double) integral of the density function over the region of density must be 1, so that

$$1 = \int_0^1 \int_0^1 K(x^2+y^2)\,dy\,dx = K \times \frac{2}{3} \;\Rightarrow\; K = \frac{3}{2}$$
$$\Rightarrow\; f(x,y) = \frac{3}{2}(x^2+y^2) \text{ for } 0 \le x \le 1 \text{ and } 0 \le y \le 1$$

In order to find the probability $P[X+Y \ge 1]$, we identify the two dimensional region representing $X+Y \ge 1$. This is generally found by drawing the boundary line for the inequality, which is $x+y=1$ (or $y=1-x$) in this case, and then determining which side of the line is represented in the inequality. We can see that $x+y \ge 1$ is equivalent to $y \ge 1-x$. This is the shaded region in the graph below.



The probability $P[X+Y \ge 1]$ is found by integrating the joint density over the two-dimensional region. It is possible to represent two-variable integrals in either order of integration. In some cases one order of integration is more convenient than the other. In this case there is not much advantage of one direction of integration over the other.

$$P[X+Y \quad 1] = \int_0^1 \int_{1-x}^1 \frac{3}{2}(x^2+y^2)\,dy\,dx = \int_0^1 \frac{1}{2}(3x^2y+y^3\Big|_{y=1-x}^{y=1})\,dx$$
$$= \int_0^1 \frac{1}{2}(3x^2+1-3x^2(1-x)-(1-x)^3)\,dx = \frac{3}{4}$$

Reversing the order of integration, we have $x \ge 1-y$, so that

$$P[X+Y \ge 1] = \int_0^1 \int_{1-y}^1 \frac{3}{2}(x^2+y^2)\,dx\,dy = \frac{3}{4}$$ □

Example 8-3:

Continuous random variables X and Y have a joint distribution with density function $f(x,y) = x^2 + \frac{xy}{3}$ for $0 < x < 1$ and $0 < y < 2$.

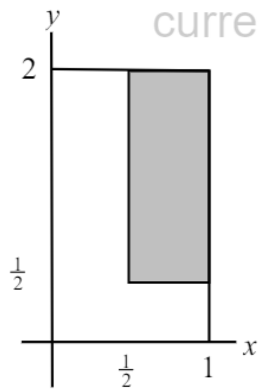
Find the conditional probability $P[X > \frac{1}{2} | Y > \frac{1}{2}]$.

Solution:

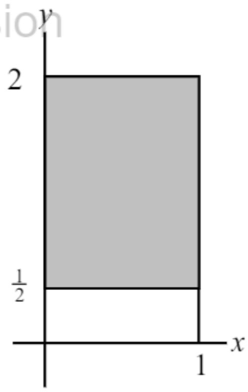
We use the usual definition $P[A|B] = \frac{P[A\cap B]}{P[B]}$ so that $P[X > \frac{1}{2} | Y > \frac{1}{2}] = \frac{P[(X>\frac{1}{2})\cap(Y>\frac{1}{2})]}{P[Y>\frac{1}{2}]}$.

These regions are described in the following diagram

For trial use only.
Content may not
represent the most
current version



$\left(X > \frac{1}{2}\right) \cap \left(Y > \frac{1}{2}\right)$



$Y > \frac{1}{2}$

$P[(X > \frac{1}{2}) \cap (Y > \frac{1}{2})] = \int_{1/2}^{1}\int_{1/2}^{2}[x^2 + \frac{xy}{3}]\,dy\,dx = \frac{43}{64}$

$P[Y > \frac{1}{2}] = \int_{1/2}^{2}[\int_0^1 f(x,y)\,dx]\,dy = \int_{1/2}^{2}\int_0^1 [x^2 + \frac{xy}{3}]\,dx\,dy = \frac{13}{16}$

$\rightarrow P[X > \frac{1}{2}|Y > \frac{1}{2}] = \frac{43/64}{13/16} = \frac{43}{52}$ □

Cumulative distribution function of a joint distribution: If random variables X and Y have a joint distribution, then the cumulative distribution function is $F(x,y) = P[(X \le x) \cap (Y \le y)]$.
In the continuous case, $F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(s,t)\,dt\,ds$
and in the discrete case, $F(x,y) = \sum_{s=-\infty}^{x}\sum_{t=-\infty}^{y} f(s,t)$. In the continuous case, $\frac{\partial^2}{\partial x\,\partial y}F(x,y) = f(x,y)$.

Example 8-4:
The cumulative distribution function for the joint distribution of the continuous random variables X and Y is $F(x,y) = 0.2 \times (3x^3y + 2x^2y^2)$ for $0 \le x \le 1$ and $0 \le y \le 1$. Find $f(\frac{1}{2}, \frac{1}{2})$.
Solution: $f(x,y) = \frac{\partial^2}{\partial x\,\partial y}F(x,y) = 0.2 \times (9x^2 + 8xy) \rightarrow f(\frac{1}{2}, \frac{1}{2}) = \frac{17}{20}$. □

Expectation of a function of jointly distributed random variables

If $h(x,y)$ is a function of two variables, and X and Y are jointly distributed random variables, then the **expected value of $h(X,Y)$** is defined to be
$\boldsymbol{E[h(X,Y)] = \sum_x\sum_y h(x,y) \times f(x,y)}$ in the discrete case, and
$\boldsymbol{E[h(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y) \times f(x,y)\,dy\,dx}$ in the continuous case.

For trial use only.
Content may not
represent the most
current version

Example 8-5:

X and Y are discrete random variables which are jointly distributed with the following probability function $f(x,y)$ (from Example 8-1):

		X		
		-1	0	1
	1	$\frac{1}{18}$	$\frac{1}{9}$	$\frac{1}{6}$
Y	0	$\frac{1}{9}$	0	$\frac{1}{6}$
	-1	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{9}$

Find $E[X \times Y]$.

Solution:

$$E[XY] = \sum_x \sum_y xy \cdot f(x,y) = (-1) \times 1 \times \frac{1}{18} + (-1) \times 0 \times \frac{1}{9} + (-1)(-1) \times \frac{1}{6}$$
$$+ 0 \times 1 \times \frac{1}{9} + 0 \times 0 \times 0 + 0 \times (-1) \times \frac{1}{9}$$
$$+ 1 \times 1 \times \frac{1}{6} + 1 \times 0 \times \frac{1}{6} + 1 \times (-1) \times \frac{1}{9} = \frac{1}{6}.$$ □

Example 8-6:

Suppose that $f(x,y) = \frac{3}{2}(x^2 + y^2)$ is the density function for the joint distribution of the continuous random variables X and Y defined over the unit square defined on the region $0 \le x \le 1$ and $0 \le y \le 1$. Find $E[X^2 + Y^2]$.

Solution:

$$E[X^2 + Y^2] = \int_0^1 \int_0^1 (x^2 + y^2) \times f(x,y)\, dy\, dx = \int_0^1 \int_0^1 (x^2 + y^2)(\tfrac{3}{2})(x^2 + y^2)\, dy\, dx$$
$$= \int_0^1 (1.5x^4 + x^2 + 0.3)\, dx = \frac{14}{15}.$$ □

Marginal distribution of X found from a joint distribution of X and Y

If X and Y have a joint distribution with joint density or probability function $f(x,y)$, then the **marginal distribution of X** has a probability function or density function denoted $f_X(x)$, which is equal to $\boldsymbol{f_X(x) = \sum_y f(x,y)}$ in the discrete case, and is equal to $\boldsymbol{f_X(x) = \int_{-\infty}^{\infty} f(x,y)\, dy}$

in the continuous case. The density function for the marginal distribution of Y is found in a similar way, $f_Y(y)$ is equal to either $f_Y(y) = \sum_x f(x,y)$ or $f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\, dx$.

For instance, $f_X(1) = \sum_y f(1,y)$ in the discrete case. What we are doing is "adding up" the probability for all points whose x-value is 1 to get the overall probability that X is 1. The marginal distribution of X describes the random behavior of X as a single random variable.

Care must be taken when the probability space is triangular or some other non-rectangular shape. In this case one must be careful to set the limits of integration properly when finding a marginal density. This is illustrated in Example 8-9 below.

If the cumulative distribution function of the joint distribution of X and Y is $F(x,y)$, then the cdf for the marginal distributions of X and Y are
$F_X(x) = \lim_{y\to\infty} F(x,y)$ and $F_Y(y) = \lim_{x\to\infty} F(x,y)$.

This concept of marginal distribution can be extended to define the marginal distribution of any one (or subcollection) variable in a multivariate distribution. Marginal probability functions and marginal density functions must satisfy all the requirements of probability and density functions. A marginal probability function must sum to 1 over all points of probability and a marginal density function must integrate to 1.

Example 8-7:

Find the marginal distributions of X and Y for the joint distribution in Example 8-1.

Solution:

The joint distribution was given as

		X		
		-1	0	1
	1	$\frac{1}{18}$	$\frac{1}{9}$	$\frac{1}{6}$
Y	0	$\frac{1}{9}$	0	$\frac{1}{6}$
	-1	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{9}$

To find the marginal probability function for X, we first note that X can be -1 , 0 or 1.
We wish to find $f_X(-1) = P[X=-1]$, $f_X(0)$ and $f_X(1)$.

As noted above, to find $f_X(x)$ we sum over the other variable Y:

$$f_X(-1) = \sum_{\text{all } y} f(-1,y) = f(-1,-1) + f(-1,0) + f(-1,1) = \tfrac{1}{6} + \tfrac{1}{9} + \tfrac{1}{18} = \tfrac{1}{3}$$

and in a similar way we get $f_X(0) = \frac{1}{9} + 0 + \frac{1}{9} = \frac{2}{9}$ and $f_X(1) = \frac{1}{9} + \frac{1}{6} + \frac{1}{6} = \frac{4}{9}$.
In Example 8-1 we saw that $P[X=0] = \frac{2}{9}$. What we were finding was the marginal probability $f_X(0)$. Note also that $\sum_{\text{all } x} f_X(x) = f_X(-1) + f_X(0) + f_X(1) = \frac{1}{3} + \frac{2}{9} + \frac{4}{9} = 1$. This verifies that $f_X(x)$ satisfies the requirements of a probability function.
The marginal probability function of Y is found in the same way, except that sum over x (across each row in the table above).
$f_Y(-1) = \frac{1}{6} + \frac{1}{9} + \frac{1}{9} = \frac{7}{18}$, $f_Y(0) = \frac{1}{9} + 0 + \frac{1}{6} = \frac{5}{18}$ and $f_Y(1) = \frac{1}{18} + \frac{1}{9} + \frac{1}{6} = \frac{1}{3}$. □

Example 8-8:

Find the marginal distributions of X and Y for the joint distribution in Example 8-2.

Solution:

The joint density function is $f(x,y) = \frac{3}{2}(x^2+y^2)$ for $0 \le x \le 1$ and $0 \le y \le 1$.

The marginal density function of X is found by integrating out the other variable y.

$f_X(x) = \int_{\text{all } y} f(x,y)\,dy = \int_0^1 f(x,y)\,dy = \int_0^1 \frac{3}{2}(x^2+y^2)\,dy = \frac{3}{2}x^2 + \frac{1}{2}$ for $0 \le x \le 1$.

We can verify that this is a proper density function by checking that $\int_0^1 f_X(x)\,dx = 1$.

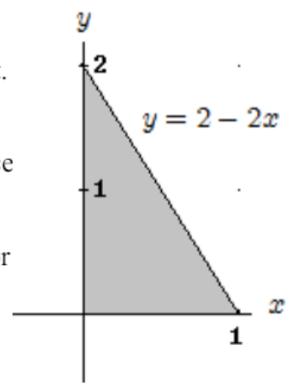
In a similar way, $f_Y(y) = \frac{3}{2}y^2 + \frac{1}{2}$ for $0 \le y \le 1$. □

Example 8-9:

Continuous random variables X and Y have a joint distribution with density function $f(x,y) = \frac{3(2-2x-y)}{2}$ in the region bounded by $y=0$, $x=0$ and $y = 2-2x$. Find the density function for the marginal distribution of X for $0 < x < 1$. Find $P[X > \frac{1}{2}]$ and find $P[X > Y]$.

Solution:

The region of joint density is illustrated in the graph at the right. Note that X must be in the interval $(0,1)$ and Y must be in the interval $(0,2)$ and the joint probability space is triangular. Since $f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$, we note that given a value of x in $(0,1)$, the possible values of y (with non-zero density for $f(x,y)$) must satisfy $0 < y < 2-2x$, so that $f_X(x) = \int_0^{2-2x} f(x,y)\,dy = \int_0^{2-2x} \frac{3(2-2x-y)}{2}\,dy = 3(1-x)^2$.

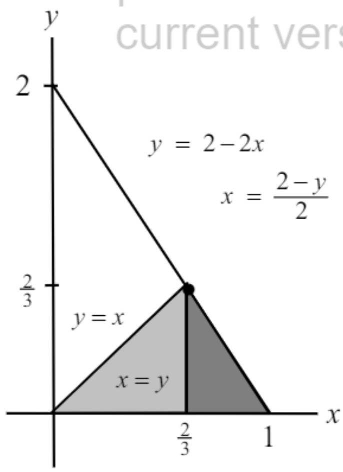


Once we have the marginal density function for X, we can find $P[X > \frac{1}{2}]$.

$P[X > \frac{1}{2}] = \int_{1/2}^1 f_X(x)\,dx = \int_{1/2}^1 3(1-x)^2\,dx = \frac{1}{8}$.

Note that we could find $P[X > \frac{1}{2}]$ by identifying the two-dimensional region and integrating $f(x,y)$. This would come out to be $\int_{1/2}^1 \int_0^{2-2x} f(x,y)\,dy\,dx = \int_{1/2}^1 f_X(x)\,dx$, which is the same as finding the marginal density first. We could also have reversed the order of integration in x and y, so that $P[X > \frac{1}{2}] = \int_0^1 \int_{1/2}^{(2-y)/2} f(x,y)\,dx\,dy$. This involves a little more algebra.

The region for which $X > Y$ is identified in the graph below. The line $y = x$ intersects with the line $y = 2-2x$ at the point $(\frac{2}{3}, \frac{2}{3})$. The region we are looking for is $y < x$, and lies below the line $y = x$.



The probability can be expressed as a double integral. If we set the order of integration with dx on the outside then we must integrate in two pieces, first from 0 to $\frac{2}{3}$ beneath the line $y = x$, and then from $\frac{2}{3}$ to 1 beneath the line $y = 2 - 2x$.

$$P[X > Y] = \int_0^{2/3}\int_0^x \frac{3(2-2x-y)}{2}\,dy\,dx + \int_{2/3}^1\int_0^{2-2x} \frac{3(2-2x-y)}{2}\,dy\,dx = \frac{8}{27} + \frac{1}{27} = \frac{1}{3}.$$

The integral can be found in the $dx\,dy$ integration order. In that case,
$P[X > Y] = \int_0^{2/3}\int_y^{(2-y)/2} \frac{3(2-2x-y)}{2}\,dx\,dy = \frac{1}{3}$.

In this case it is more efficient to express the integral with dy on the outside.
There is one other note on this example. The probability space was originally described as "the region bounded by $y = 0$, $x = 0$ and $y = 2 - 2x$". We might also see this region defined in the following way: $0 < x < \frac{2-y}{2}$ and $y > 0$. The reader can check that this is the same region. □

Independence of random variables X and Y

Random variables X and Y with density functions $f_X(x)$ and $f_Y(y)$ are said to be independent (or stochastically independent) if the probability space is rectangular ($a \le x \le b$, $c \le y \le d$, where the endpoints can be infinite) and if the joint density function is of the form $\boldsymbol{f(x,y) = f_X(x) \times f_Y(y)}$. Independence of X and Y is also equivalent to the factorization of the cumulative distribution function $\boldsymbol{F(x,y) = F_X(x) \times F_Y(y)}$ **for all** $\boldsymbol{(x,y)}$.

For the discrete joint distribution in Example 8-1 we can see that X and Y are not independent, because, for instance, $f(-1,-1) = \frac{1}{6} \ne \frac{1}{3} \times \frac{7}{18} = f_X(-1) \times f_Y(-1)$. For the continuous joint distribution of Example 8-8, we see that

$$f(x,y) = \frac{3}{2}(x^2 + y^2) \ne (\frac{3}{2}x^2 + \frac{1}{2})(\frac{3}{2}y^2 + \frac{1}{2}) = f_X(x) \times f_Y(y)\,,$$

so X and Y are not independent.

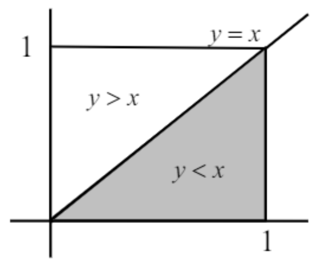
Example 8-10:

Suppose that X and Y are independent continuous random variables with the following density functions: $f_X(x) = 1$ for $0 < x < 1$ and $f_Y(y) = 2y$ for $0 < y < 1$. Find $P[Y < X]$.

Solution:

Since X and Y are independent, the density function of the joint distribution of X and Y is $f(x,y) = f_X(x) \times f_Y(y) = 2y$, and is defined on the rectangle created by the intervals for X and Y, which, in this case, is the unit square. The graph below illustrates the region for the probability in question.

$P[Y < X] = \int_0^1 \int_0^x 2y\,dy\,dx = \frac{1}{3}$.



□

Conditional distribution of Y given $X = x$

The way in which a conditional distribution is defined follows the basic definition of conditional probability, $P[A|B] = \frac{P[A \cap B]}{P[B]}$. In fact, given a discrete joint distribution, this is exactly how a conditional distribution is defined. Example 8-1 described a discrete joint distribution of X and Y, and then Example 8-7 showed how to formulate the marginal distributions of X and Y. We now wish to formulate a conditional distribution. For instance, for the joint distribution of Example 8-1, suppose we wish to describe the conditional distribution of X given $Y = 1$. What we are trying to describe are conditional probabilities of the form $P[X = x|Y = 1]$.

We find these conditional probabilities in the usual way that conditional probability is defined.

$$P[X = -1|Y = 1] = \frac{P[(X=-1) \cap (Y=1)]}{P[Y=1]}$$

The denominator is the marginal probability that $Y = 1$, $f_Y(1) = \frac{1}{3}$. The numerator is the joint probability $f(-1,1) = \frac{1}{18}$, which is found in the joint probability table. Then,

$P[X = -1|Y = 1] = \frac{f(-1,1)}{f_Y(1)} = \frac{1/18}{1/3} = \frac{1}{6}$. We would denote this conditional probability $f_{X|Y}(-1|Y = 1)$. In a similar way, we can get $f_{X|Y}(0|Y = 1) = \frac{f(0,1)}{f_Y(1)} = \frac{1/9}{1/3} = \frac{1}{3}$, and $f_{X|Y}(1|Y = 1) = \frac{f(1,1)}{f_Y(1)} = \frac{1/6}{1/3} = \frac{1}{2}$. This completely describes the conditional distribution of X given $Y = 1$. As with any discrete distribution, probabilities for a conditional distribution must add to 1, and this is the case for this conditional distribution, since

$f_{X|Y}(-1|Y = 1) + f_{X|Y}(0|Y = 1) + f_{X|Y}(1|Y = 1) = \frac{1}{6} + \frac{1}{3} + \frac{1}{2} = 1$.

A conditional distribution satisfies all the same properties of any distribution. We can find a conditional mean, a conditional variance, etc. For instance, the conditional mean of X given $Y=1$ in the example we have just been considering is

$$\begin{aligned} E[X|Y=1] &= \sum_{all\, x} x \times f_{X|Y}(x|Y=1) \\ &= (-1) \times f_{X|Y}(-1|Y=1) + 0 \times f_{X|Y}(0|Y=1) + 1 \times f_{X|Y}(1|Y=1) \\ &= (-1) \times \tfrac{1}{6} + 0 \times \tfrac{1}{3} + 1 \times \tfrac{1}{2} = \tfrac{1}{3}. \end{aligned}$$

We can find the second conditional moment in a similar way,

$$E[X^2|Y=1] = \sum_{all\, x} x^2 \times f_{X|Y}(x|Y=1)$$

Then the conditional variance would be

$$Var[X|Y=-1] = E[X^2|Y=-1] - (E[X|Y=-1])^2$$

The expression for conditional probability that was used above in the discrete case was $\boldsymbol{f_{X|Y}(x|Y=y) = \frac{f(x,y)}{f_Y(y)}}$. This can be applied to find a conditional distribution of Y given $X=x$ also, so that we define $f_{Y|X}(y|X=x) = \frac{f(x,y)}{f_X(x)}$.

We also apply this same algebraic form to define the conditional density in the continuous case, with $f(x,y)$ being the joint density and $f_X(x)$ being the marginal density. In the continuous case, the conditional mean of Y given $X=x$ would be

$$\boldsymbol{E[Y|X=x] = \int y \times f_{Y|X}(y|X=x)\, dy}$$

where the integral is taken over the appropriate interval for the conditional distribution of Y given $X=x$. The conditional density/probability is also written as $f_{Y|X}(y|x)$, or $f(y|x)$.

If X and Y are independent random variables, then

$$f_{Y|X}(y|X=x) = \frac{f(x,y)}{f_X(x)} = \frac{f_X(x) \cdot f_Y(y)}{f_X(x)} = f_Y(y)$$

and in a similar way we have $f_{X|Y}(x|Y=y) = f_X(x)$, which indicates that the density of Y does not depend on X and vice-versa.

The conditional density function must satisfy the usual requirement of a density function, $\int_{-\infty}^{\infty} f_{Y|X}(y|x)\, dy = 1$. Note also that if $f_X(x)$, the marginal density of X is known, and if $f_{Y|X}(y|X=x)$, the conditional density of Y given $X=x$, is also known, then the joint density of X and Y can be formulated as

$$\boldsymbol{f(x,y) = f_{Y|X}(y|X=x) \times f_X(x)}$$

Example 8-11:

Find the conditional distribution of Y given $X = -1$ for the joint distribution of Example 8-1. Find the conditional expectation of Y given $X = -1$.

Solution:

The marginal probability function for X was found in Example 8-7, where it was found that $f_X(-1) = \frac{1}{3}$. The conditional probability function of Y given $X = -1$ is

$f_{Y|X}(y|X = -1) = \frac{f(-1,y)}{f_X(-1)} = \frac{f(-1,y)}{1/3}$. Then

$f_{Y|X}(-1|X = -1) = \frac{f(-1,-1)}{1/3} = \frac{1/6}{1/3} = \frac{1}{2}$, $f_{Y|X}(0|X = -1) = \frac{f(-1,0)}{1/3} = \frac{1/9}{1/3} = \frac{1}{3}$,

and $f_{Y|X}(1|X = -1) = \frac{f(-1,1)}{1/3} = \frac{1/18}{1/3} = \frac{1}{6}$.

$E[Y|X = -1] = \sum_{\text{all } y} y \cdot f_{Y|X}(y|X = -1) = (-1) \times \frac{1}{2} + 0 \times \frac{1}{3} + 1 \times \frac{1}{6} = -\frac{1}{3}$. □

Example 8-12:

Find the conditional density and conditional expectation and conditional variance of X given $Y = .3$ for the joint distribution of Example 8-2.

Solution:

$f_{X|Y}(x|Y = 0.3) = \frac{f(x,0.3)}{f_Y(0.3)} = \frac{\frac{3}{2}(x^2+(0.3)^2)}{\frac{3}{2}(0.3)^2+\frac{1}{2}} = \frac{\frac{3}{2}(x^2+0.09)}{0.635}$.

The conditional expectation is

$E[X|Y = .3] = \int_0^1 x \times f_{X|Y}(x|Y = 0.3)\, dx = \int_0^1 x \times \frac{\frac{3}{2}(x^2+0.09)}{0.635}\, dx = 0.697$.

The conditional second moment of X given $Y = 0.3$ is

$E[X^2|Y = 0.3] = \int_0^1 x^2 \times f_{X|Y}(x|Y = 0.3)\, dx = \int_0^1 x^2 \times \frac{\frac{3}{2}(x^2+0.09)}{0.635}\, dx = 0.543$.

The conditional variance is

$Var[X|Y = 0.3] = E[X^2|Y = 0.3] - (E[X|Y = 0.3])^2 = 0.543 - (0.697)^2 = 0.057$. □

Example 8-13:

Continuous random variables X and Y have a joint distribution with density function

$f(x,y) = \frac{\pi}{2}(\sin\frac{\pi}{2}y)e^{-x}$ for $0 < x < \infty$ and $0 < y < 1$. Find $P[X > 1|Y = \frac{1}{2}]$.

Solution:

$P[X > 1|Y = \frac{1}{2}] = \frac{P[(X>1)\cap(Y=\frac{1}{2})]}{f_Y(\frac{1}{2})}$

$P[(X > 1) \cap (Y = \frac{1}{2})] = \int_1^\infty f(x,\frac{1}{2})\, dx = \int_1^\infty \frac{\pi}{2}(\sin\frac{\pi}{4})e^{-x}dx = \frac{\pi\sqrt{2}}{4} \times e^{-1}$

$f_Y(\frac{1}{2}) = \int_0^\infty f(x,\frac{1}{2})\, dx = \int_0^\infty \frac{\pi}{2}(\sin\frac{\pi}{4})e^{-x}dx = \frac{\pi\sqrt{2}}{4} \Rightarrow P[X > 1|Y = \frac{1}{2}] = e^{-1}$

Note that $f_X(x) = \int_0^1 f(x,y)\, dy = \int_0^1 \frac{\pi}{2}(\sin\frac{\pi}{2}y)e^{-x}\, dy = e^{-x}$ and

$f_Y(y) = \int_0^\infty f(x,y)\, dx = \int_0^\infty \frac{\pi}{2}(\sin\frac{\pi}{2}y)e^{-x}\, dx = \frac{\pi}{2}(\sin\frac{\pi}{2}y)$.

Then we see that X and Y are independent.

This follows from $f(x,y) = (e^{-x})(\frac{\pi}{2} \times sin\,\frac{\pi}{2}y) = f_X(x) \times f_Y(y)$.
From independence it follows that $P[X > 1|Y = \frac{1}{2}] = P[X > 1] = e^{-1}$ (same as the result above).

As a final comment on this example, we could have noticed at the start that since $f(x,y)$ can be factored into separate functions in x and y, we might anticipate that X and Y are independent. Since the joint distribution is defined on a rectangular area, it follows that X and Y must be independent if the pdf factors into a function of x only multiplied by a function of y only.

Furthermore, the factor involving x is e^{-x}. Since this integrates to 1 over the range for X, it must be the marginal density of X, $f_X(x) = e^{-x}$. Then, because of independence, we have $P[X > 1|Y = \frac{1}{2}] = P[X > 1]$, and from the marginal density, this is $\int_1^{\infty} e^{-x}\,dx = e^{-1}$. □

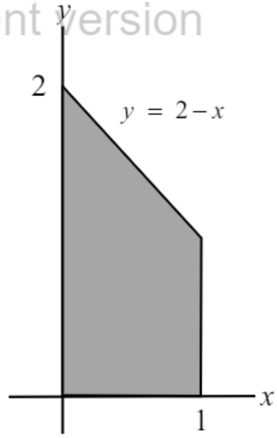
Example 8-14:
X is a continuous random variable with density function $f_X(x) = x + \frac{1}{2}$ for $0 < x < 1$. X is also jointly distributed with the continuous random variable Y, and the conditional density function of Y given $X = x$ is $f_{Y|X}(y|X = x) = \frac{x+y}{x+\frac{1}{2}}$ for $0 < x < 1$ and $0 < y < 1$.
Find $f_Y(y)$ for $0 < y < 1$.

Solution:
$f(x,y) = f(y|x) \cdot f_X(x) = \frac{x+y}{x+\frac{1}{2}} \times (x + \frac{1}{2}) = x + y$
Then, $f_Y(y) = \int_0^1 f(x,y)\,dx = y + \frac{1}{2}$. □

As Example 8-14 shows, we can construct the joint density $f(x,y)$ from knowing the conditional density $f_{Y|X}(y|x)$ and the marginal density $f_X(x)$ using the relationship $f(x,y) = f(y|x) \times f_X(x)$. When doing this, care must be taken to ensure that proper two-dimensional region is being formulated for the joint distribution. The following example illustrates this point.

Example 8-15:
A company will experience a loss X that is uniformly distributed between 0 and 1. The company pays a bonus to its employees that is uniformly distributed on the interval $(0, 2 - X)$, which depends on the amount of the loss that occurred. Find the expected amount of the bonus paid.

Solution:
X has marginal pdf $f_X(x) = 1$ for $0 < x < 1$. Let Y be the bonus paid. The conditional distribution of Y given $X = x$ is a uniform distribution on the interval $(0, 2 - x)$, with conditional density $f_{Y|X}(y|x) = \frac{1}{2-x}$. The joint density of X and Y is $f(x,y) = f_{Y|X}(y|x) \cdot f_X(x) = 1 \times \frac{1}{2-x} = \frac{1}{2-x}$ and the two-dimensional region of probability is the region $0 < y < 2 - x$ and $0 < x < 1$. This region is the shaded area in the graph at the right.



The expected value of Y can be found as $E[Y]=\int_0^1\int_0^{2-x} y\times\frac{1}{2-x}\,dy\,dx=\frac{3}{4}$.
We could have found the marginal density on Y, and then have found $E[Y]$, but that is more awkward because for $0<y\leq 1$ we have $f_Y(y)=\int_0^1\frac{1}{2-x}\,dx=ln\,2$, and for $1<y\leq 2$ we have $f_Y(y)=\int_0^{2-y}\frac{1}{2-x}\,dx=ln\,2\,-\,ln\,y$. □

Example 8-16:

Suppose that X has a continuous distribution with pdf $f_X(x)=2x$ on the interval $(0,1)$, and $f_X(x)=0$, elsewhere. Suppose that Y is a continuous random variable such that the conditional distribution of Y given $X=x$ is uniform on the interval $(0,x)$. Find the mean and variance of Y.

Solution:

We find the unconditional (marginal) distribution of Y. We are given $f_X(x)=2x$ for $0<x<1$, and $f_{Y|X}(y|X=x)=\frac{1}{x}$ for $0<y<x$. Then, $f(x,y)=f(y|x)\cdot f_X(x)=\frac{1}{x}\cdot 2x=2$ for $0<y<x<1$.
The unconditional (marginal) distribution of Y has pdf.
$f_Y(y)=\int_{-\infty}^{\infty}f(x,y)\,dx=\int_y^1 2\,dx=2(1-y)$ for $0<y<1$ (and $f_Y(y)$ is 0 elsewhere). Then $E[Y]=\int_0^1 y\cdot 2(1-y)\,dy=\frac{1}{3}$, $E[Y^2]=\int_0^1 y^2\cdot 2(1-y)\,dy=\frac{1}{6}$,
and $Var[Y]=E[Y^2]-(E[Y])^2=\frac{1}{6}-(\frac{1}{3})^2=\frac{1}{18}$. □

Covariance between random variables X and Y: If random variables X and Y are jointly distributed with joint density/probability function $f(x,y)$, X and Y might not be independent of one another. We have seen some examples in which X and Y are independent, and some in which they are not. There are a couple of numerical measures that describe, in some sense, the dependence that exists between X and Y. Covariance is one such measure.

The covariance between X and Y is

$$\boldsymbol{Cov[X,Y]=E[(X-E[X])(Y-E[Y])]}$$
$$\boldsymbol{=E[(X-\mu_X)(Y-\mu_Y)]=E[XY]-E[X]\times E[Y]}$$

A positive covariance between X and Y is an indication that "large" values of X (values of X that are bigger than $E[X]$) tend to occur paired with "large" values of Y, and the same for "small" values of X and Y. Negative covariance indicates the opposite relationship. Covariance near or at 0 indicates that the "size" of X value is not related to the "size" of the Y values to which they are paired.

Note that $Cov[X,X] = Var[X]$. One important application of the covariance is in finding the variance of the sum of X and Y. Suppose that a, b and c are constants. Then it can be shown that

$$\boldsymbol{Var[aX + bY + c] = a^2 Var[X] + b^2 Var[Y] + 2ab Cov[X,Y]}$$

Coefficient of correlation between random variables X and Y:

The coefficient of correlation between random variables X and Y is
$\boldsymbol{\rho(X,Y) = \rho_{X,Y} = \frac{Cov[X,Y]}{\sigma_X \sigma_Y}}$, where σ_X and σ_Y are the standard deviations of X and Y respectively. Note that $-1 \leq \rho_{X,Y} \leq 1$ always.

Example 8-17:

Find $Cov[X,Y]$ for the jointly distributed discrete random variables in Example 8-1 above.

Solution:

$Cov[X,Y] = E[XY] - E[X] \times E[Y]$. In Example 8-5 it was found that $E[XY] = \frac{1}{6}$. The marginal probability function for X is $P[X=1] = \frac{1}{6} + \frac{1}{6} + \frac{1}{9} = \frac{4}{9}$,
$P[X=0] = \frac{2}{9}$ and $P[X=-1] = \frac{1}{3}$, and the mean of X is
$E[X] = 1 \times \frac{4}{9} + 0 \times \frac{2}{9} + (-1) \times \frac{1}{3} = \frac{1}{9}$.

In a similar way, the probability function of Y is found to be $P[Y=1] = \frac{1}{3}$, $P[Y=0] = \frac{5}{18}$, and
$P[Y=-1] = \frac{7}{18}$, with a mean of $E[Y] = -\frac{1}{18}$.
Then, $Cov[X,Y] = \frac{1}{6} - \frac{1}{9} \times (-\frac{1}{18}) = \frac{14}{81}$. □

Example 8-18: The coefficient of correlation between random variables X and Y is $\frac{1}{3}$, and $\sigma_X^2 = a$, $\sigma_Y^2 = 4a$. The random variable Z is defined to be $Z = 3X - 4Y$, and it is found that $\sigma_Z^2 = 114$. Find a.

Solution:

$\sigma_Z^2 = Var[Z] = 9Var[X] + 16Var[Y] + 2 \cdot (3)(-4)\,Cov[X,Y]$.
Since $Cov[X,Y] = \rho[X,Y] \times \sigma_X \times \sigma_Y = \frac{1}{3} \times \sqrt{a} \times \sqrt{4a} = \frac{2a}{3}$, it follows that
$114 = \sigma_Z^2 = 9a + 16 \times 4a - 24 \times \frac{2a}{3} = 57a \rightarrow a = 2$. □

Moment generating function of a joint distribution: Given jointly distributed random variables X and Y, the moment generating function of the joint distribution is
$M_{X,Y}(t_1,t_2) = E[e^{t_1X+t_2Y}]$. This definition can be extended to the joint distribution of any number of random variables. It can be shown that $E[X^nY^m]$ is equal to the multiple partial derivative evaluated at 0, $E[X^nY^m] = \frac{\partial^{n+m}}{\partial^n t_1\, \partial^m t_2}\, M_{X,Y}(t_1,t_2)\Big|_{t_1=t_2=0}$.

Example 8-19:

Suppose that X and Y are random variables whose joint distribution has moment generating function $M(t_1,t_2) = (\frac{1}{4}e^{t_1} + \frac{3}{8}e^{t_2} + \frac{3}{8})^{10}$, for all real t_1 and t_2.
Find the covariance between X and Y.

Solution:

$$Cov[X,Y] = E[XY] - E[X]\cdot E[Y]$$

$$E[XY] \;=\; \frac{\partial^2}{\partial t_1\,\partial t_2}\, M_{X,Y}(t_1,t_2)\Big|_{t_1=t_2=0}$$

$$= 10\times 9\times (\tfrac{1}{4}e^{t_1} + \tfrac{3}{8}e^{t_2} + \tfrac{3}{8})^8(\tfrac{1}{4}e^{t_1})(\tfrac{3}{8}e^{t_2})\Big|_{t_1=t_2=0} = \frac{135}{16}$$

$$E[X] = \frac{\partial}{\partial t_1}\, M_{X,Y}(t_1,t_2)\Big|_{t_1=t_2=0} = 10\times(\tfrac{1}{4}e^{t_1} + \tfrac{3}{8}e^{t_2} + \tfrac{3}{8})^9(\tfrac{1}{4}e^{t_1})\Big|_{t_1=t_2=0} = \frac{5}{2}$$

$$E[Y] = \frac{\partial}{\partial t_2}\, M_{X,Y}(t_1,t_2)\Big|_{t_1=t_2=0} = 10\times(\tfrac{1}{4}e^{t_1} + \tfrac{3}{8}e^{t_2} + \tfrac{3}{8})^9(\tfrac{3}{8}e^{t_2})\Big|_{t_1=t_2=0} = \frac{15}{4}$$

$$\Rightarrow \quad Cov[X,Y] = \frac{135}{16} - \frac{5}{2}\times\frac{15}{4} = -\,\frac{15}{16}.$$ □

The Bivariate normal distribution:

Suppose that X and Y are normal random variables with means and variances
$E[X] = \mu_X$, $Var[X] = \sigma_X^2$, $E[Y] = \mu_Y$, $Var[Y] = \sigma_Y^2$, and with correlation coefficient ρ_{XY}.
X and Y are said to have a bivariate normal distribution. The conditional mean and variance of Y given $X = x$ are

$$\boldsymbol{E[Y|X=x] = \mu_Y + \rho_{XY}\times\frac{\sigma_Y}{\sigma_X}\cdot(x-\mu_X) = \mu_Y + \frac{Cov(X,Y)}{\sigma_X^2}\times(x-\mu_X)}$$

and

$$\boldsymbol{Var[Y|X=x] = \sigma_Y^2\times(1-\rho_{XY}^2)}\,.$$

Similar relationships apply for the conditional distribution of X given $Y = y$.
If normal random variables X and Y are independent, then $\rho_{XY} = 0$, and vice-versa.
The pdf of the bivariate normal distribution is

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}\times exp\Big[-\frac{1}{2(1-\rho^2)}\times\Big[\big(\tfrac{x-\mu_X}{\sigma_X}\big)^2 + \big(\tfrac{y-\mu_Y}{\sigma_Y}\big)^2 - 2\rho\big(\tfrac{x-\mu_X}{\sigma_X}\big)\big(\tfrac{y-\mu_Y}{\sigma_Y}\big)\Big]\Big]\,.$$

The bivariate normal has occurred infrequently on Exam P (there is one question on this topic in the Exam P sample questions on the SOA website).

Some results and formulas related to this section are

(i) If X and Y are independent, then for any functions g and h,
$E[g(X) \cdot h(Y)] = E[g(X)] \times E[h(Y)]$, and in particular, $E[X \times Y] = E[X] \times E[Y]$

(ii) The density/probability function of jointly distributed variables X and Y can be written in the form
$f(x, y) = f_{Y|X}(y|X = x) \times f_X(x) = f_{X|Y}(x|Y = y) \times f_Y(y)$

(iii) $Cov[X, Y] = E[X \times Y] - \mu_X \times \mu_Y = E[XY] - E[X] \times E[Y]$
$Cov[X, Y] = Cov[Y, X]$. If X and Y are independent, then $E[X \times Y] = E[X] \times E[Y]$
and $Cov[X, Y] = 0$. For constants a, b, c, d, e, f
and random variables X, Y, Z and W,
$Cov[aX + bY + c\ ,\ dZ + eW + f]$
$= adCov[X, Z] + aeCov[X, W] + bdCov[Y, Z] + beCov[Y, W]$

(iv) $Var[X + Y] = E[(X + Y)^2] - (E[X + Y])^2$
$= E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2$
$= E[X^2] + E[2XY] + E[Y^2] - (E[X])^2 - 2E[X]E[Y] - (E[Y])^2$
$= Var[X] + Var[Y] + 2 \cdot Cov[X, Y]$

If X and Y are independent, then $Var[X + Y] = Var[X] + Var[Y]$.
For any X, Y, $Var[aX + bY + c] = a^2Var[X] + b^2Var[Y] + 2ab \times Cov[X, Y]$

(v) **If X and Y have a joint distribution which is uniform (constant density) on the two dimensional region R (usually R will be a triangle, rectangle or circle in the (x, y) plane) then the pdf of the joint distribution is $\frac{1}{Area\,of\,R}$ inside the region R (and the pdf is 0 outside). The probability of any event A (represented by a subset of R) is the proportion $\frac{Area\,of\,A}{Area\,of\,R}$. Also the conditional distribution of Y given $X = x$ has a uniform distribution on the line segment (or segments) defined by the intersection of the region R with the line $X = x$.**
The marginal distribution of Y might or might not be uniform

(vi) $E[h_1(X, Y) + h_2(X, Y)] = E[h_1(X, Y)] + E[h_2(X, Y)]$, and in particular,
$E[X + Y] = E[X] + E[Y]$ and $E[\sum X_i] = \sum E[X_i]$

(vii) $\lim_{x \to -\infty} F(x, y) = \lim_{y \to -\infty} F(x, y) = 0$