ST5201: Basic Statistical Theory
Chapter 3: Joint Distributions

Choi, Yunjin
stachoiy@nus.edu.sg

Department of Statistics and Applied Probability
National University of Singapore (NUS)

September 5th, 2017

Outline

- Introduction
- Joint Distribution of Discrete Random Variables
- Joint Distribution of Continuous Random Variables
- Independent Random Variables
- Conditional Distributions
- Functions of Jointly Distributed Random Variables
- Extrema and Order Statistics

Learning Outcomes

- Questions to Address: How to understand multiple r.v.s simultaneously ⋆ What are marginal distributions/densities & joint distribution/density ⋆ What are independent r.v.'s ⋆ What a conditional distribution is ⋆ How to characterize functions of multiple r.v.'s ⋆ What are order statistics

Concept & Terminology

- joint/marginal/conditional distribution $\star$ joint probability $\star$ multinomial distribution $\star$ bivariate normal distribution
- joint cumulative distribution function $\star$ joint/marginal/conditional probability mass function $\star$ joint/marginal/conditional probability density function $\star$ univariate/multivariate distribution
- independence $\star$ Jacobian $\star$ sums of r.v.s $\star$ Order statistics $\star$ Median

Mandatory Reading

Textbook: Section 3.1 – Section 3.7

Motivating Examples

- In the population of NUS students (sample space Ω), we may be interested in the following characteristics of a student, his/her gender (G), major degree (M), and year of studies (Y). Here, for each student, (G, M, Y) denotes a student's gender, major degree, and year of studies. Each of these 3 characteristics can be represented as a r.v.
- Any investor is interested in the return of his/her investments in a portfolio that contains several stocks & funds in both local & overseas markets
  - Random returns of the individual derivatives: $X_1, X_2, \cdots, X_N$
  - For diversification: How $X_1, X_2, \cdots, X_N$ vary simultaneously?
  - The total return is $Y = v(X_1, X_2, \cdots, X_N) = X_1 + \cdots + X_N$

- Up to Chapter 2, we study only 1 r.v. at a time
- Obviously, it is possible to define many r.v.'s from one single experiment through different rule of association. This is a common situation.
  New-born baby examination: weight, height, ...
- Usually, we are interested in ≥ 2 r.v.'s defined from the same Ω simultaneously. We call the joint probability structure of ≥ 2 r.v.'s as joint distribution:
  - bivariate distribution of $X \& Y$;
  - multivariate distribution of $X_1, X_2, \cdots, X_n$;
  - For only 1 r.v., we have a univariate distribution.

Recall: The pmf of a discrete r.v. X specifies how much probability mass is assigned to each possible value of X, with $p(x_i) = P(X = x_i)$.

Definition

Let X & Y be 2 discrete r.v.'s defined on the same sample space Ω. The *joint frequency function* or *joint probability mass function*, $p(x, y)$, is defined for each pair $(x, y) \in \mathbb{R} \times \mathbb{R}$ by

$$p(x, y) = P(X = x, Y = y).$$

- $0 \le p(x, y) \le 1$, $\sum_x \sum_y p(x, y) = 1$
- For the RHS of the formula
  - Re-write $\{X = x\} = \{\omega \in \Omega | X(\omega) = x\}$ & $\{Y = y\} = \{\omega \in \Omega | Y(\omega) = y\}$
  - Get all the common ω's in the 2 sets, that $\{X = x\} \cap \{Y = y\}$
  - Summation of the probs of these ω's, that $P(\{X = x\} \cap \{Y = y\})$

- Prob that (X, Y) is in any set C in the xy-plane (e.g., $C = \{(x, y) \in \mathbb{R}^2 | x + y = 5\}$ or $C = \{(x, y) \in \mathbb{R}^2 | x/y \leq 3\}$)

Probability of Any Event

Let $C \subset \mathbb{R}^2$ be any set consisting of (x, y) values. Then,

$$\underline{P((X, Y) \in C)} = \sum_{(x,y)\in C} p(x, y).$$

Marginal pmf

The *marginal prob mass functions* of X&Y, denoted by p_X & p_Y, can be recovered from the joint pmf as

$$p_X(x) = P(X = x) = \sum_y p(x, y), \quad \&$$
$$p_Y(y) = P(Y = y) = \sum_x p(x, y).$$

Toss 3 fair coins:

- $\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$
- Define the r.v.'s $\begin{cases} X = \#\text{ of tails in the first 2 tosses} \\ Y = \#\text{ of heads in all the 3 tosses} \end{cases}$. Clearly, X taks on $0, 1, 2$, and Y takes on 0, 1, 2, 3.
- Mapping from all outcomes $\omega \in \Omega$ to the values of X and Y

ω	HHH	HHT	HTH	THH	HTT	THT	TTH	TTT
$X(\omega)$	0	0	1	1	1	1	2	2
$Y(\omega)$	3	2	2	2	1	1	1	0

- Hence, we have possible values for (X, Y) as $(0, 3), (0, 2), (1, 2), (1, 1), (2, 1), (2, 0)$, and

$$\begin{aligned} p(0,3) &= P(X = 0, Y = 3) = P(\{HHH\}) = 1/8 \\ p(1,1) &= P(X = 1, Y = 1) = P(\{HTT, THT\}) = 2/8 \end{aligned}$$

- Similarly to the calculation before, we have $p(0,3) = 1/8$, $p(0,2) = 1/8$, $p(1,2) = 2/8$, $p(1,1) = 2/8$, $p(2,1) = 1/8$, $p(2,0) = 1/8$, $p(x,y) = 0$, otherwise.

- The joint pmf $p(x,y)$ of X & Y in a tabular form:

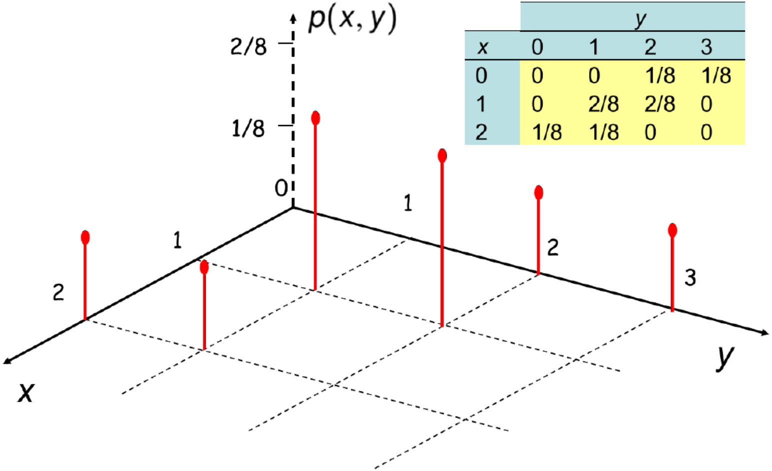
	y			
x	0	1	2	3
0	0	0	1/8	1/8
1	0	2/8	2/8	0
2	1/8	1/8	0	0

- With the joint pmf, probability of events related to 2 r.v.'s can be found. For example, $P(X < 2, Y \geq 1) = 3/4$ & $P(X = Y) = 1/4$

- Summing up the columns or rows in the above table gives p_X or p_Y

Example: 2 Discrete r.v.'s

- Similar to the graph of the pmf of 1 discrete r.v., we can also sketch the joint pmf



	y			
x	0	1	2	3
0	0	0	1/8	1/8
1	0	2/8	2/8	0
2	1/8	1/8	0	0

Example: Trinomial Distribution I

The binomial dist. can be generalized to multinomial distribution, where there are K possible outcomes.
When $K = 3$, we call it as a trinomial distribution:

- Suppose that there is an experiment with 3 possible outcomes (say, success, failure, "undecided") in place of a Bernoulli trial
- Probs of 3 outcomes: p_1, p_2, $p_3 = 1 - p_1 - p_2$
- Repeat independently the experiment for n times
- Define $X = \#$ of successes, $Y = \#$ of failures, $Z = n - X - Y = \#$ of "undecided"

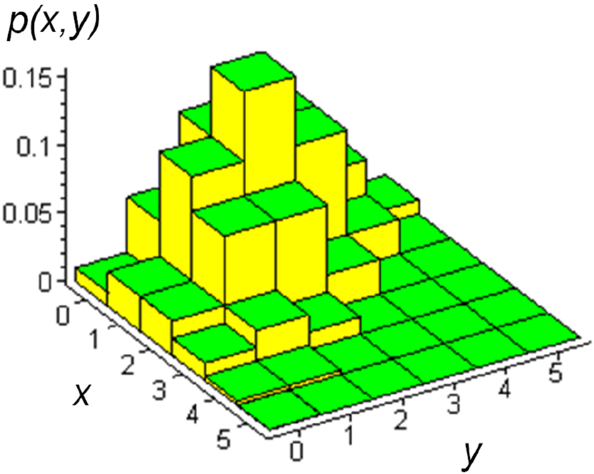
The pair (X, Y) is said to have a *trinomial distribution with # of trials n and probs of success p_1, p_2*, and its joint pmf is given by

$$p(x, y) = \binom{n}{x \quad y \quad n-x-y} p_1^x p_2^y (1 - p_1 - p_2)^{n-x-y}$$

for $x, y = 0, 1, \cdots n$, s.t. $x + y \leq n$, and 0 otherwise.

Trinomial Dsitribution II

A 3-d plot of the pmf of a trinomial distribution with # of trials $n = 5$ & probs of success $p_1 = 1/5$, $p_2 = 2/5$:



The *marginal distributions* for X, Y and Z are binomial, i.e., $X \sim Bin(n, p_1)$, $Y \sim Bin(n, p_2)$, $Z \sim Bin(n, p_3)$.

Example for $p_X(x)$: Sum up all the possible values of Y given that $X = x$ is of interest. When $X = x$, Y can take on values from 0 up to $n - x$. For $x = 0, 1, \cdots, n$,

$$\begin{aligned}
p_X(x) &= \sum_y p(x,y) = \sum_{y=0}^{n-x} \binom{n}{x \quad y \quad n-x-y} p_1^x p_2^y p_3^{n-x-y} \\
&= \frac{n!}{x!} p_1^x \sum_{y=0}^{n-x} \frac{1}{y!(n-x-y)!} p_2^y p_3^{n-x-y} \\
&= \frac{n!}{x!(n-x)!} p_1^x (1-p_1)^{n-x} \sum_{y=0}^{n-x} \frac{(n-x)!}{(n-x-y)!} \left(\frac{p_2}{1-p_1}\right)^y \left(\frac{p_3}{1-p_1}\right)^{n-x-y}
\end{aligned}$$

The sum in the latter expression is 1 as it is the sum of all probs of a $Bin(n-x, \frac{p_2}{p_2+p_3})$ r.v.

For $m \geq 3$ discrete r.v.'s, $X_1, \cdots, X_m$, defined on the same Ω:

- joint pmf: $p_{X_1,\cdots,X_m}(x_1, \cdots, x_m) = P(X_1 = x, \cdots, X_m = x_m)$
- *1-dimensional (1-dim) marginal pmf*'s:

$$p_{X_i}(x_i) = \sum_{x_1,\cdots,x_{i-1},x_{i+1},x_m} p_{X_1,\cdots X_m}(x_1, \cdots, x_m)$$

  for all i
- *2-dim joint pmf*'s: e.g.

$$p_{X_1,X_2}(x_1, x_2) = \sum_{x_3,\cdots,x_m} p_{X_1,X_2,\cdots,X_m}(x_1, x_2, \cdots, x_m)$$

- *3-dim*, $\cdots$, *$m-1$-dim joint pmf*'s

Joint Cumulative Distribution Function I

Recall: The cdf is a universal way to characterize all kinds of r.v.'s (either discrete or cont.)

Definition

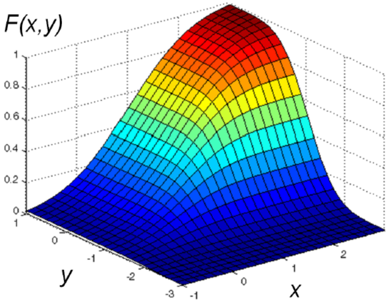
The joint random behaviour of any 2 r.v.'s, X & Y, is determined by the *joint cumulative distribution function (joint cdf)*:

$$F(x, y) = P(X \leq x, Y \leq y) \qquad -\infty < x, y < \infty$$

- $F(\cdot, \cdot)$ is a function from $\mathbb{R}^2 \equiv \mathbb{R} \times \mathbb{R}$ to $[0, 1]$
- Sometimes written as $F_{X,Y}(x, y)$ for clarity
- $\{X \leq x, Y \leq y\} \equiv \{X \leq x\} \cap \{Y \leq y\}$: similar to the joint pmf

Joint Cumulative Distribution Function II

- Instead of a 2-d graph for plotting the cdf of 1 r.v., the joint cdf F is plotted visually as a 3−d graph with
  - X-axis: values of X
  - Y-axis: values of Y
  - Z-axis (i.e., the vertical axis): values of $F(x, y)$ for different x & y
- An example plot of a joint cdf of 2 cont. r.v.'s:



Properties of $F(x, y)$

- $0 \leq F(x, y) \leq 1$
- Non-decreasing function in both x & y:

$$\begin{array}{ll} \text{if } a < c, \text{ then } F(a, y) \leq F(c, y), & -\infty < y < \infty \\ \text{if } b < c, \text{ then } F(x, b) \leq F(x, c), & -\infty < x < \infty \end{array}$$

- $F(-\infty, -\infty) = F(-\infty, y) = F(x, -\infty) = 0, \; F(\infty, \infty) = 1$
- cdf's(also called *marginal pdf's*) of X & of Y:

$$F_X(x) = F(x, \infty) \qquad \& \qquad F_Y(y) = F(\infty, y)$$

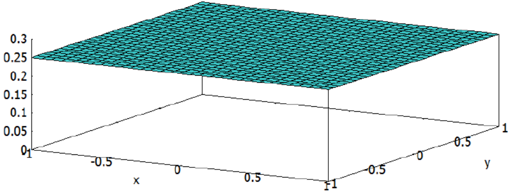
Definition

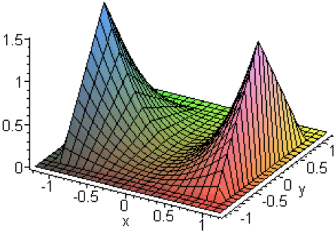
The *joint (prob) density function (joint pdf)* of 2 cont. r.v.'s X & Y is an integrable function $f : \mathbb{R}^2 \to [0, \infty)$ satisfying

- $f(x, y) > 0$ for any $(x, y) \in \mathbb{R}^2$
- $f(x, y)$ is a piecewise cont. function of x & of y
- $P((X, Y) \in \mathbb{R}^2) = 1$
- The prob that (X, Y) takes on a pair (x, y) in any set $C \subset \mathbb{R}^2$ equals the volume of the object/space over the region C and bounded by the surface $z = f(x, y)$, which is expressible as a double integral:

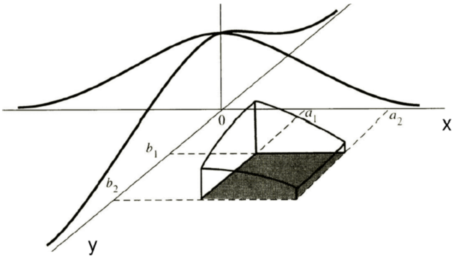
$$P((X, Y) \in C) = \int_C \int f(x, y) dy\, dx = \int_C \int f(x, y) dx\, dy$$

- The simplest example of a joint pdf, f, is a constant function, which is equivalent to a flat surface on a bounded set $S \subset \mathbb{R}^2$ (plotted as below on the left), where S is called *the support of f*
- Plots of 2 joint pdf's $f(x, y)$ of 2 cont. r.v.'s (at the right: f is not a constant function)





- Ideally:
  - Support $S = \mathbb{R}^2$
  - Set of interest $C = \{(x, y) \in \mathbb{R}^2 : a_1 \le x \le a_2, b_1 \le y \le b_2\}$, where a_1, a_2, b_1, b_2 are some fixed constants, is the rectangle in black in the xy-plane below



Following properties of joint cdf,

$$P((X, Y) \in C) = P(a_1 \le X \le a_2, b_1 \le Y \le b_2) = \int_{a_1}^{a_2} \int_{b_1}^{b_2} f(x, y)dy\,dx$$

- Commonly, $S \subset \mathbb{R}^2$, i.e., only on a set S, $f(x,y) = g(x,y) \neq 0$. Following properties of joint cdf, $P((X,Y) \in C)$ equals

$$\begin{aligned}
\int_{C\cap\mathbb{R}^2}\int f(x,y)dy\,dx &= \int_{C\cap(S\cup(\mathbb{R}^2\setminus S))}\int f(x,y)dy\,dx \\
&= \int_{C\cap S}\int f(x,y)dy\,dx + \int_{C\cap(\mathbb{R}^2\setminus S)}\int f(x,y)dy\,dx \\
&= \underline{\int_{C\cap S}\int f(x,y)dy\,dx} + \int_{C\cap(\mathbb{R}^2\setminus S)}\int 0dy\,dx
\end{aligned}$$

$$P((X,Y) \in C) = \int_{C\cap S}\int f(x,y)dy\,dx$$

Computing Joint Probability III

Evaluating a Double Integral

For any set $D \subset \mathbb{R}^2$ & a known function $g(x, y) \neq 0$, a *double integral*

$$\int_D \int g(x, y) dy\, dx \tag{1}$$

is evaluated by

- integrating over all possible values of y in the inner integral holding x (the variable in the outer integral) as a constant at each of its possible value
- integrating over all possible values of x in the outer integral

- **Our goal**: Find some expressions, a_1, a_2, b_1, b_2, as the limits such as (1) becomes

$$\int_D \int g(x,y)dy\,dx = \int_{a_1}^{a_2} \int_{b_1}^{b_2} g(x,y)dy\,dx, \qquad (2)$$

so that we can evaluate the inner integral in y by treating x as a constant & then the outer integral in x
Note: b_1, b_2 can be functions of x

- **In practice**, *plot the domain of integration D* on the xy-plane to

  (a) locate all possible values of x as $a_1 \leq x \leq a_2$ from D
  (b) locate all possible values of y as $b_1 \leq y \leq b_2$ from D with *x fixed at each of its possible value on $[a_1, a_2]$*

  If (b) fails, break the interval (a_1, a_2) for x in (a) into 2 or more pieces so that the RHS of (2) becomes a sum of 2 or more double integrals

Example: Joint pdf & Joint Prob, I

Consider

$$f(x,y) = \begin{cases} 24xy, & 0 \le x \le 1, 0 \le y \le 1, x+y \le 1 \\ 0, & otherwise \end{cases}$$

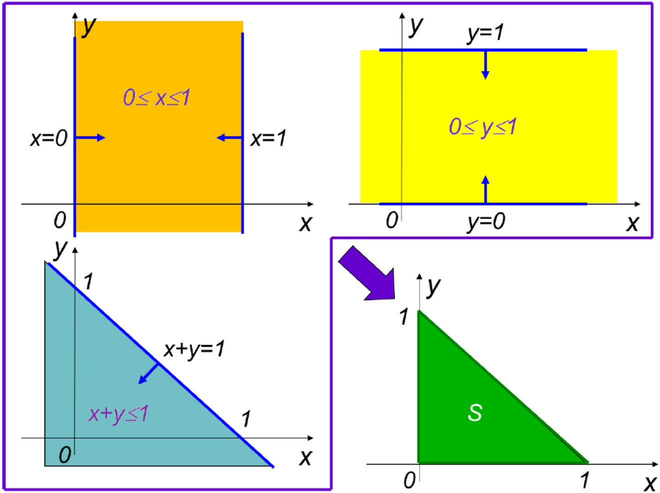
- We can show that it is a proper joint pdf as follows:
  - Clearly, $f(x,y) \ge 0$ for $-\infty < x, y < \infty$, & f is piecewise cont.
  - Let $S = \{(x,y) \in \mathbb{R}^2 | 0 \le x \le 1, 0 \le y \le 1, x+y \le 1\}$
  - To verify that the double integral of f over $\mathbb{R}^2$ is 1, it involves computing the double integral

$$\begin{aligned} \int_{\mathbb{R}^2}\int f(x,y)dy\,dx &= \int_S\int f(x,y)dy\,dx + \int_{\mathbb{R}^2\setminus S}\int f(x,y)dy\,dx \\ &= \int_S\int 24xydy\,dx + \int_{\mathbb{R}^2\setminus S}\int 0dy\,dx \end{aligned}$$

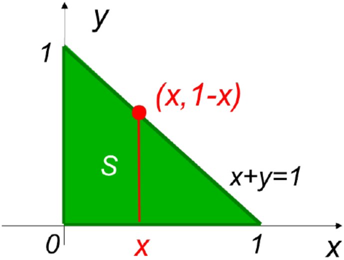
Refer to (1), we have $g(x,y) = 24xy$ & $D = S$

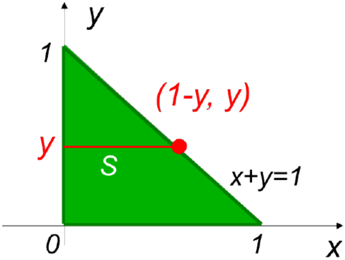
Next, draw the domain of integration S

Example: Joint pdf & Joint Prob, II



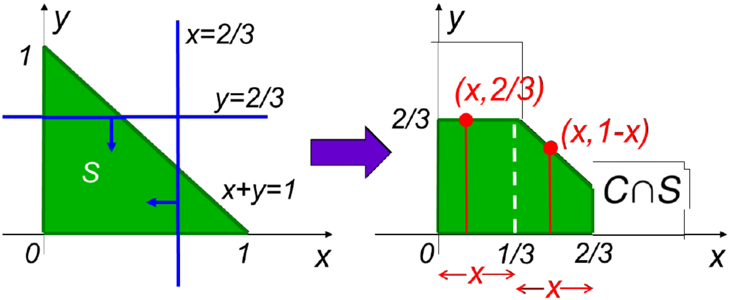
$$\int_S\int 24xy\,dy\,dx = 24\int_0^1\left[\int_0^{1-x} xy\,dy\right]dx = 24\int_0^1 x\left[\frac{y^2}{2}\right]_0^{1-x} dx$$
$$= 12\int_0^1 x(1-x)^2\,dx = 12\int_0^1 (x-2x^2+x^3)\,dx = 1$$





Alternatively, $\int_S\int 24xy\,dx\,dy = 24\int_0^1\left[\int_0^{1-y} xy\,dx\right]dy = 1$

2. Now, suppose that we are interested in *computing the prob that both X & Y take on values less than* $2/3$, *i.e.*, $P((X,Y) \in C)$ where $C = \left\{(x,y) \in \mathbb{R}^2 : x < 2/3, y < 2/3\right\}$
   - From (2), it suffices to compute the *double integral (3)* with $g(x,y) = 24xy$ & $D = C \cap S$ represented by the **green pentagon** below at the right



Example: Joint pdf & Joint Prob, V

- Unfortunately, it is impossible to assign unique limits for the inner integral when we consider all x values from 0 to $2/3$, so we split $C \cap S$ into into 2 pieces wrt x from 0 to $1/3$ & from $1/3$ to $2/3$ in order to obtain unique limits for the inner integrals over the 2 resulting regions

$$
\begin{aligned}
P(X < 2/3, Y < 2/3) &= \int_{C\cap S}\int 24xy dy\, dx \\
&= \int_0^{1/3}\int_0^{2/3} 24xy dy\, dx + \int_{1/3}^{2/3}\int_0^{1-x} 24xy dy\, dx \\
&= \int_0^{1/3} \frac{16}{3} x dx + \int_{1/3}^{2/3}\int_0^{1-x} 12x(1-x)^2 dx \\
&= \frac{8}{27} + [6x^2 - 8x^3 + 3x^4]|_{1/3}^{2/3} = 7/9
\end{aligned}
$$

Marginal cdf's & Marginal pdf's

Marginal Cumulative Distribution Functions

The *marginal cdf*'s of X & of Y, denoted by F_X & F_Y, defined as

$$F_X(x) = F(x, \infty), \qquad \& \qquad F_Y(y) = F(\infty, y).$$

Marginal Probability Density Functions or Marginal Densities

The *marginal pdf*'s of X & of Y, denoted by f_X & f_Y, are given by

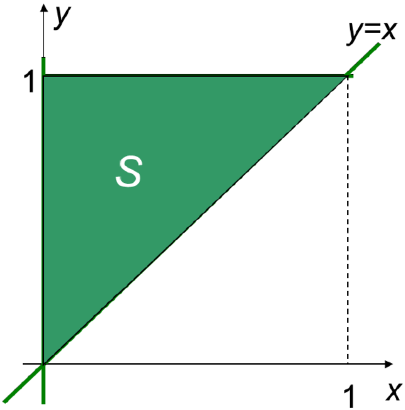
$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy, \qquad \& \qquad f_Y(y) = \int_{-\infty}^{\infty} f(x, y)dx.$$

- Note: when $S \subset \mathbb{R}^2$, the latter 2 integrals become
  1. $\int_{b_1}^{b_2} g(x, y)dy$: $[b_1, b_2]$ gives all possible values of y when x is a fixed value in S
  2. $\int_{a_1}^{a_2} g(x, y)dx$: $[a_1, a_2]$ gives all possible values of x when y is a fixed value in S

Suppose that the joint pdf of X & Y is defined by

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

1. What is $P\left(X \le \frac{3}{4}, Y \ge \frac{1}{2}\right)$?
2. Find the marginal pdf's



To compute the required prob, draw the *domain of integration* $C \cap S \equiv A \cup B$ of the double integral $\int_{C\cap S} 6(1-y)\,dy\,dx$ as in the graph below (where $C = \{(x,y) \in \mathbb{R} | x < 3/4, y > 1/2\}$):

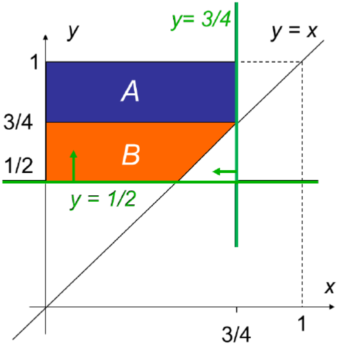
$$\int_A \int 6(1-y)\,dx\,dy + \int_B \int 6(1-y)\,dx\,dy$$

$$= \int_{3/4}^{1} \int_0^{3/4} 6(1-y)\,dx\,dy$$

$$+ \int_{1/2}^{3/4} \int_0^{y} 6(1-y)\,dx\,dy$$

$$= \int_{3/4}^{1} \frac{9}{2}(1-y)dy + \int_{1/2}^{3/4} 6(y-y^2)dy$$

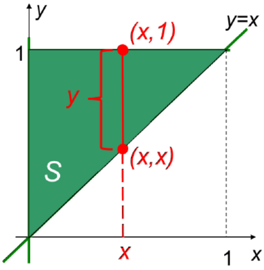
$$= \frac{9}{2}\left[y - \frac{y^2}{2}\right]_{3/4}^{1} + \left[3y^2 - 2y^3\right]_{1/2}^{3/4} = \frac{31}{64}$$



For $0 < x < 1$ (*i.e.,* all possible values of X from the support S),

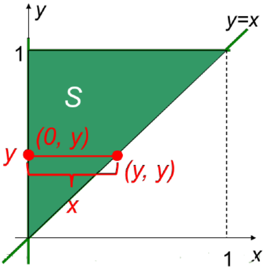
$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_x^1 6(1-y)\,dy$$
$$= \left[6y - 3y^2\right]_1^x = 3(1-x)^2$$

Note: Fix x at a certain value beforehand



For $0 < y < 1$ (*i.e.,* all possible values of Y from the support S),

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx = \int_0^y 6(1-y)\,dx$$
$$= 6y(1-y)$$



Example: Bivariate Normal Distribution

Generalize the normal distribution to ≥ 2 r.v.'s, for which we call *multivariate normal/multinormal distribution*. It is the most importnat and commonly used joint distribution of ≥ 2 r.v.'s.

Definition

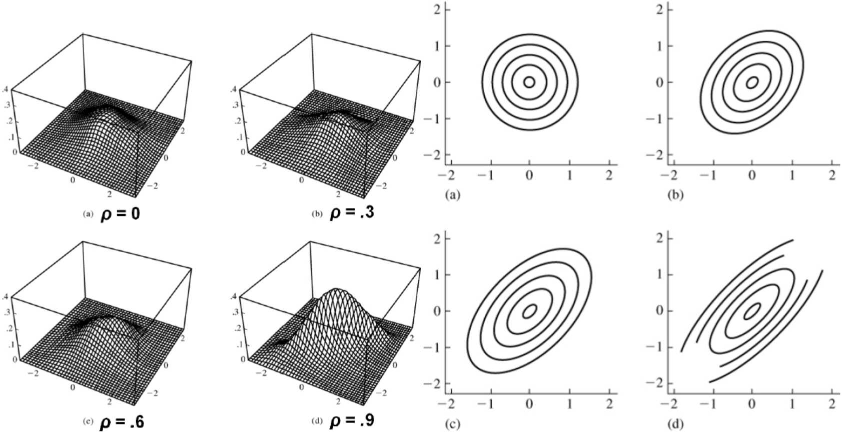
A random vector (X, Y) has a *bivariate normal distribution* with parameters $\mu_X, \mu_Y \in \mathbb{R}$, $\sigma_X, \sigma_Y > 0$, $-1 < \rho < 1$, if its joint pdf is given by, for all $x, y \in \mathbb{R}$,

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp[-\frac{1}{2(1-\rho^2)}(z_x^2 + z_y^2 - 2\rho z_x z_y)],$$

where $z_x = \frac{x-\mu_X}{\sigma_X}$, $z_y = \frac{y-\mu_Y}{\sigma_Y}$.

- $\frac{x-\mu_X}{\sigma_X}$ has the same form as standardization of normal r.v.
- 5 parameters for the *family of bivariate normal distributions*
- $\exp(x) = e^x$ is usual exponential function in an exponential density

With $\mu_X = \mu_Y = 0$ & $\sigma_X = \sigma_Y = 1$:



Marginal Distributions of a Bivariate Normal Distribution

For a bivariate normally distributed vector (X, Y), the *marginal distributions of X & of Y are both normal*

The marginal pdf of X is $f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy$, for $-\infty < x < \infty$.
Marking the change of variables $v = z_y = (y - \mu_Y)/\sigma_Y$ & rewriting $u = z_x$ give

$$f_X(x) = \frac{1}{2\pi\sigma_X\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \exp[-\frac{1}{2(1-\rho^2)}(u^2 + v^2 - 2\rho uv)]dv.$$

Apply the technique of completing the square with

$$u^2 + v^2 - 2\rho uv = (v - \rho u)^2 + u^2(1 - \rho^2)$$

& recognize that the resulting integral has integrand as the pdf of a $N(\rho u, 1 - \rho^2)$ r.v.
Hence, the RHS reduces to the pdf of a $N(\mu_X, \sigma_X^2)$ r.v. Analogously, $Y \sim N(\mu_Y, \sigma_Y^2)$

Example: Marginal pdf's

The marginal pdf's of $f(x,y) = \frac{3}{2}x^2(1-|y|)$, $-1 < x < 1$, $-1 < y < 1$, can be found as follows

Solution:

$$\begin{aligned}
f_X(x) &= \int_{-1}^{1} \frac{3}{2}x^2(1-|y|)dy = \frac{3}{2}x^2[\int_0^1 (1-y)dy + \int_{-1}^{0}(1+y)dy] \\
&= \frac{3}{2}x^2([y - \frac{y^2}{2}]|_0^1 + [y + \frac{y^2}{2}]|_{-1}^0) \\
&= \frac{3}{2}x^2(1 - 1/2 + 1 - 1/2) = \frac{3}{2}x^2, \qquad -1 < x < 1 \\
f_Y(y) &= \frac{3}{2}(1-|y|)\int_{-1}^{1} x^2 dx = 1 - |y|, \qquad -1 < y < 1
\end{aligned}$$

Note: One can then compute any prob statement about only X or only Y based on f_X or f_Y, & the marginal cdf's, F_X & F_Y.

Consider another joint pdf defined on the same support S, where $S = \{0 \le x \le 1, 0 \le y \le 1, x + y \le 1\}$. $f(x, y) = cxy^2$, for $(x, y) \in S$, and 0 otherwise. Find the marginal pdf's of X & of Y, & deduce the value of c.

Solution:

1. The marginal pdf of X is defined by $f_X(x) = \int_{\mathbb{R}} f(x, y)dy$ for any possible value x of X (i.e., from 0 to 1 based on the support S of f), i.e., for $0 \le 0 \le 1$,

$$f_X(x) = \int_S cxy^2 dy = c\int_0^{1-x} xy^2 dy = \frac{c}{3}x[y^3]|_0^{1-x} = \frac{c}{3}x(1-x)^3$$

2. Similarly, the marginal pdf of Y is given by, for $0 \le y \le 1$,

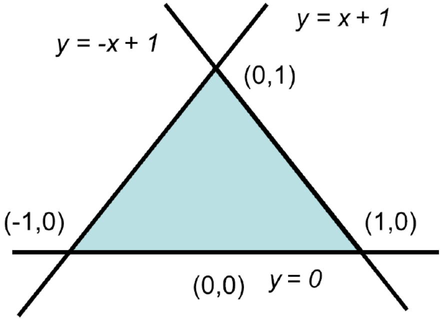
$$f_Y(y) = \int_S cxy^2 dx = c\int_0^{1-y} xy^2 dx = \frac{c}{2}y^2[x^2]|_0^{1-y} = \frac{c}{2}y^2(1-y)^2$$

3. $c = 60$ by matching the above pdf's with the pdf's of a $B(2, 4)$ r.v. & of a $B(3, 3)$ r.v.; in fact $X \sim B(2, 4)$ & $Y \sim B(3, 3)$

Example: Marginal pdf & Joint Prob I

Suppose that X & Y are jointly *uniformly distributed* over a triangle Δ defined below, *i.e.*, for some constant $c > 0$,

$$f(x,y) = \begin{cases} c, & (x,y) \in \Delta \\ 0, & \text{otherwise} \end{cases}$$



Find

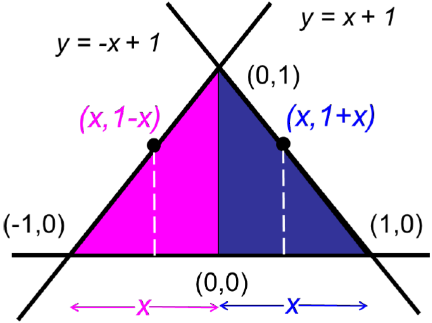
1. the value of c
2. $f_X(x)$
3. $P\left(X < \frac{3}{4}, Y < \frac{3}{4}\right)$

Solution:

1. Equating the volume of a triangular cylinder with height c & base given by the shaded triangle gives $c = 1$

Example: Marginal pdf & Joint Prob II

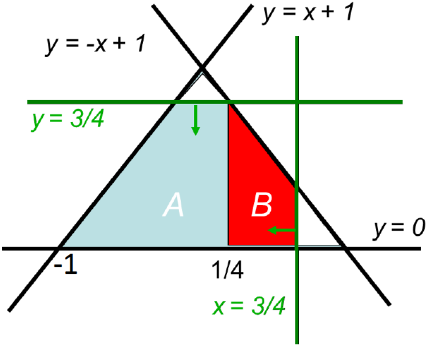
2. The graph below illustrates that *possible values of Y take different forms* when $-1 < x < 0$ & $0 < x < 1$ $\Rightarrow$ *Consider the pdf of X in 2 cases* corresponding to the 2 regions shown below



$$f_X(x) = \begin{cases} \int_0^{1+x} 1\, dy = 1 + x, & -1 < x < 0 \\ \int_0^{1-x} 1\, dy = 1 - x, & 0 < x < 1 \\ 0, & x \leq -1 \text{ or } x \geq 1 \end{cases}$$

3. Insert the lines $x = 3/4$ & $y = 3/4$ to get $C \cap S$ $\Rightarrow$ *The required prob* equals the volume of a cylinder with height as 1 & base defined by $C \cap S \equiv A \cup B$, *i.e.*,

$$P\left(X < \frac{3}{4}, Y < \frac{3}{4}\right)$$
$$= \int_A\int dx\,dy + \int_B\int dx\,dy$$
$$= \int_0^{3/4}\int_{y-1}^{1/4} 1\,dx\,dy$$
$$+ \int_{1/4}^{3/4}\int_0^{1-x} 1\,dy\,dx$$
$$= \frac{29}{32}$$



For $m \geq 3$ cont. r.v.'s, $X_1, X_2, \ldots, X_m$, defined on the same Ω:

- Joint pdf: $f_{X_1,\cdots,X_m}(x_1,\cdots,x_m) = \frac{\partial^m}{\partial x_1 \cdots \partial x_m} F_{X_1,\cdots,X_m}(x_1,\cdots,x_m)$
- *1-dim marginal pdf*'s expressible as $(m-1)$-folded integrals:

$$f_{X_i}(x_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1,\cdots,X_m}(x_1,\cdots,x_m) dx_1 \cdots dx_{i-1}\, dx_{i+1} \cdots dx_m$$

  for all i

- *2-dim joint pdf*'s expressible as $(m-2)$-folded integrals: e.g.,

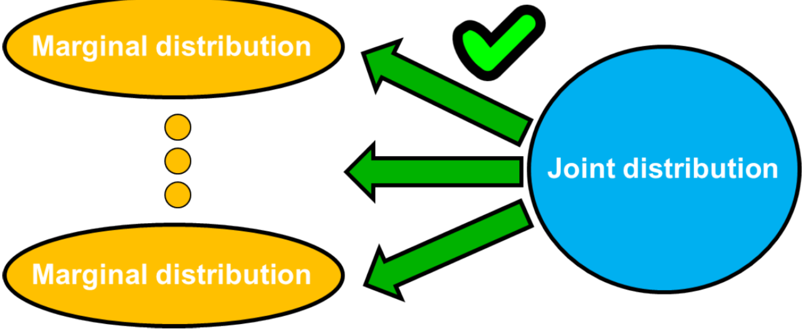
$$f_{X_1,X_2}(x_1,x_2) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1,\cdots,X_m}(x_1,\cdots,x_m) dx_3 \cdots dx_m$$

- *3-dim*, ..., *$(m-1)$-dim joint pdf*'s similarly defined
- Joint prob for any $C \subset \mathbb{R}^m$ expressible as an m-folded integral:

$$P((X_1,\cdots,X_m) \in C) = \int_C \cdots \int f_{X_1,\cdots,X_m}(x_1,\cdots,x_m) dx_1 \cdots dx_m$$

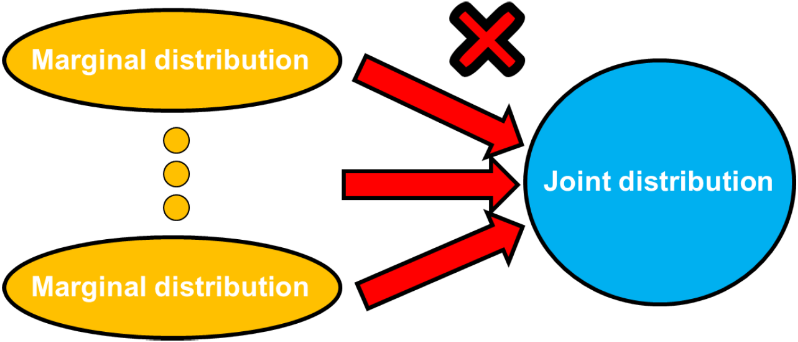
Joint Distribution Gives All 1-dim Marginal Distributions

All 1-dim marginal distributions/densities can always be obtained from the joint distributions/densities



A precautionary Note About Marginal & Joint Distributions

In general, knowing all 1-dim marginal distributions/densities do not determine the joint distribution/density



Suppose that (X_1, Y_1) is bivariate normally distributed with parameters

$$\mu_{X_1} = 0, \mu_{Y_1} = -1, \sigma_{X_1} = 1, \sigma_{Y_1} = 2, \rho = .9$$

& (X_2, Y_2) is bivariate normally distributed with parameters

$$\mu_{X_2} = 0, \mu_{Y_2} = -1, \sigma_{X_2} = 1, \sigma_{Y_2} = 2, \rho = -.3$$

- Only μ_X, μ_Y, σ_X, Σ_Y, but not ρ, appear in the marginal distributions of X & of Y
  $\Rightarrow X_1, X_2 \sim N(0, 1)$ & $Y_1, Y_2 \sim N(-1, 2^2)$, i.e., (X_1, Y_1) & (X_2, Y_2) share the same marginal distributions
- Given only these 2 marginal distributions, $N(0, 1)$& $N(-1, 2^2)$, we would never know the value of ρ, &, in turn, the joint distribution

There exist exceptions of the result discussed at the previous 2 pages: Knowing all the 1-dim marginal distributions/densities determine the joint distribution/density when all the r.v.'s are indept

Definition

Random variables $X_1, \ldots, X_n$ are said to be *independent (indept)* if their joint cdf factors into the product of their marginal cdf's

$$F(x_1, \cdots, x_n) = F_{X_1}(x_1) \times F_{X_2}(x_2) \times \cdots \times F_{X_n}(x_n)$$

for all values $x_1, \cdots, x_n \in \mathbb{R}$; otherwise $X_1, \cdots, X_n$ are said to be *dependent*.

- *Alternative definitions*: replace cdf's by pmf's & pdf's for discrete & cont. r.v.'s, respectively

Independent Random Variables II

Definition

Random variables $X_1, \ldots, X_n$ are said to be *independent (indept)* if their joint pmf factors into the product of their marginal pmf's

$$p(x_1, \cdots, x_n) = p_{X_1}(x_1) \times p_{X_2}(x_2) \times \cdots \times p_{X_n}(x_n)$$

for all values $x_1, \cdots, x_n \in \mathbb{R}$.

Definition

Random variables $X_1, \ldots, X_n$ are said to be *independent (indept)* if their joint pdf factors into the product of their marginal pdf's

$$f(x_1, \cdots, x_n) = f_{X_1}(x_1) \times f_{X_2}(x_2) \times \cdots \times f_{X_n}(x_n)$$

for all values $x_1, \cdots, x_n \in \mathbb{R}$.

- These three definitions are equivalent.

Checking For Independence

- All the 3 previous definitions about indep of r.v.'s require all the $n+1$ expressions/functions at both the LHS & the RHS of the displayed identities (in theory, given the joint cdf/pmf/pdf at the LHS, one can obtain the n marginal expressions at the RHS by *marginalizations*, but the computations could be very tedious & costly in time)

A Shortcut for Checking Independence

Random variables $X_1, \cdots, X_n$ are indept if and only if (iff) there exist functions $g_1, \cdots, g_n \colon \mathbb{R} \to \mathbb{R}$ s.t. for all $x_1, \cdots, x_n \in \mathbb{R}$, we have the joint pmf/pdf

$$f(x_1, \cdots, x_n) = g_1(x_1) \times g_2(x_2) \times \cdots \times g_n(x_n)$$

- *To check for indep*: Re-arrange the expression of the joint pmf/pdf to see if it is a product of some terms, each of which involves at most a single x_i

Some Properties About Indept. r.v.'s

When r.v.'s $X_1, \cdots, X_n$ are indept,

(a) Joint prob equals ta product of the marginal probs:

$$P(X_1 \in A_1, \cdots, X_n \in A_n) = P(X_1 \in A_1) \times \cdots \times P(X_n \in A_n)$$

for any set $A_1, \cdots, A_n \subset \mathbb{R}$

(b) For functions $h_1, \cdots, h_n$: $\mathbb{R} \to \mathbb{R}$, define r.v.'s $Y_i = h_i(X_i)$ for $i = 1, 2, \cdots, n$, then

$$Y_1, \cdots, Y_n \text{ are indept}$$

Example: Verify Independence

1. Consider the example on Page 9. X & Y are dependent as

$$0 = P(X=0, Y=0) \neq P(X=0)P(Y=0) = \frac{1}{8}\frac{2}{8} = \frac{1}{32}$$

2. On Page 24, X & Y jointly distributed with pdf
$f(x,y) = \begin{cases} 24xy, & 0 \leq x \leq 1, 0 \leq y \leq 1, x+y \leq 1 \\ 0, & \text{otherwise} \end{cases}$, are dependent
as values of X & Y are related through $x+y \leq 1$ in the support S

3. For the bivariate normal distribution defined in Example at Page 34, only when $\rho = 0$, X & Y are indept as the joint pdf factors into a product of 2 terms involving only x & only y, respectively:

$$f(x,y) = \overbrace{\left\{\frac{1}{2\pi\sigma_X\sigma_Y}\exp\left[-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right]\right\}}^{\text{in terms of } x \text{ only}} \times \overbrace{\left\{\exp\left[-\frac{1}{2}\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]\right\}}^{\text{in terms of } y \text{ only}}$$

for $x, y \in \mathbb{R}$

Example: Verify Independence By Shortcut

Suppose that

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y}, & 0 < x < \infty, 0 < y < \infty \\ 0, & \text{otherwise} \end{cases}$$

Let

$$g(x) = \begin{cases} 2e^{-x}, & 0 < x < \infty \\ 0, & \text{otherwise} \end{cases} \quad \& \quad h(y) = \begin{cases} e^{-2y}, & 0 < y < \infty \\ 0, & \text{otherwise} \end{cases}$$

Then, it is easily verified that

$$f(x,y) = g(x)h(y), \qquad x, y \in \mathbb{R}$$

$\therefore$ X & Y are *indept*

- Suppose that X & Y are both discrete, & they are indept. Then, one can obtain the joint pmf

$$p(x,y) = p_X(x)p_Y(y), \qquad \text{for all } x, y$$

by multiplying the marginal pmf's when they are known

- Suppose that X & Y are indept standard normal r.v.'s. The joint pdf of X & Y is then given by

$$\begin{aligned} f(x,y) &= f_X(x)f_Y(y) \\ &= \frac{1}{\sqrt{2\pi}}e^{-x^2/2} \cdot \frac{1}{\sqrt{2\pi}}e^{-y^2/2} \\ &= \frac{1}{2\pi}e^{-(x^2+y^2)/2}, \qquad x, y \in \mathbb{R} \end{aligned}$$

Conditional Distributions: Discrete Case I

Definition

Given the joint pmf $p_{X,Y}$ of 2 discrete r.v.'s X & Y, if $p_Y(y) > 0$ for some y, the *conditional prob mass function of X given that $Y = y$* is defined by

$$\begin{aligned} P_{X|Y}(x|y) &= P(X = x|Y = y) \\ &= \frac{P(X = x, Y = y)}{P(Y = y)} \\ &= \frac{p_{X,Y}(x, y)}{p_Y(y)} \end{aligned}$$

for all values $x \in \mathbb{R}$.

- A proper pmf in x:
  - $0 \leq p_{X|Y}(x|y) \leq 1$ for all x
  - $\sum_x p_{X|Y}(x|y) = 1$

Conditional Distributions: Discrete Case II

The Idea: We are interested in an event $A = \{X \text{ takes on the value } x\}$ knowing that another event $B = \{Y \text{ takes on the value } y\}$ has happened, & can follow the same idea & discussion about conditional prob as in Chap. 1

- In analogy to discussion about conditional prob in Chap. 1, the phrase "given that $Y = y$" or the notation "$Y|y$" refers to that Y is known to take on the value y, i.e.,

  Only outcomes $\omega \in \Omega$ s.t. $Y(\omega) = y$ are possible to occur

- **Note**: Only 1 r.v. X is random but Y is not random anymore because it has already happened to take on the value y & y is a constant throughout

Conditional Distributions: Discrete Case III

Definition

$X|Y = y$ is called a *conditional r.v.* defined on the reduced sample space consisting of only outcomes $\omega \in \Omega$ s.t. $Y(\omega) = y$ instead of the original sample space Ω.

- Just another discrete r.v. as those discussed before, which is characterized by the conditional pmf $p_{X|Y}(x|y)$
- Take on values from the collection of all possible values of $X \Rightarrow$ Take on at most the same # of values as X
- In general, different from X
- Same as X only when X & Y are indept: $p_{X|Y}(x|y) = p_X(x)$ for all values $x, y \in \mathbb{R}$

Conditional Distributions: Discrete Case IV

Treat $A = \{X = x\}$ & $B = \{Y = y\}$ again, & follow the discussion in Chap. 1 about multiplication law & law of total prob:

Multiplication Law & Law of Total Prob.

For any value $x, y \in \mathbb{R}$

- *Multiplication law*: The joint pmf can be re-expressed as

$$p_{X,Y}(x, y) = p_Y(y)p_{X|Y}(x|y) = p_X(x)p_{Y|X}(y|x)$$

- *Law of total prob.*: The marginal pmf's equal

$$p_X(x) = \sum_y p_Y(y)p_{X|Y}(x|y) \qquad (3)$$

$$p_Y(y) = \sum_x p_X(x)p_{Y|X}(y|x) \qquad (4)$$

Consider Example at Page 9. We can define 7 possible & distinct conditional r.v.'s:

$$X|Y = 0,\ X|Y = 1,\ X|Y = 2,\ X|Y = 3,\ Y|X = 0,\ Y|X = 1,\ Y|X = 2$$

e.g., the conditional r.v. $X|Y = 1$ is defined by

$$p_{X|Y}(1|1) = \frac{2/8}{3/8} = \frac{2}{3}$$

$$p_{X|Y}(2|1) = \frac{1/8}{3/8} = \frac{1}{3}$$

$$p_{X|Y}(x|1) = 0, \qquad x \neq 1, 2$$

	y			
x	0	1	2	3
0	0	0	1/8	1/8
1	0	2/8	2/8	0
2	1/8	1/8	0	0
p(y)	1/8	3/8	3/8	1/8

If X & Y are indept $Poi(\lambda)$ & $Poi(\mu)$ r.v.'s, the conditional pmf of X given that $Z = X + Y = n > 0$, for any $j = 0, 1, \ldots, n$, is defined by

$$
\begin{aligned}
P(X = j|Z = n) &= \frac{P(X = j,\ X + Y = n)}{P(X + Y = n)} = \frac{P(X = j,\ Y = n - j)}{P(X + Y = n)} \\
&= \frac{P(X = j)P(Y = n - j)}{P(X + Y = n)} = \frac{\dfrac{e^{-\lambda}\lambda^j}{j!} \times \dfrac{e^{-\mu}\mu^{n-j}}{(n-j)!}}{\dfrac{e^{-(\lambda+\mu)}(\lambda + \mu)^n}{n!}} \\
&= \binom{n}{j}\left(\frac{\lambda}{\lambda + \mu}\right)^j \left(1 - \frac{\lambda}{\lambda + \mu}\right)^{n-j}
\end{aligned}
$$

where the 3rd & the 4th equalities follow from indep of X & Y, & that $Z = X + Y$ is a $Poi(\lambda + \mu)$ r.v. (to be shown), respectively

$\therefore X|Z = n \sim Bin\left(n, \dfrac{\lambda}{\lambda + \mu}\right)$

Consider Example at Page 12. The conditional pmf of $X|Y = y$, for $y = 0, 1 \ldots, n$, is

$$p_{X|Y}(x|y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

$$= \frac{\binom{n}{x\ y\ n-x-y} p_1^x p_2^y (1 - p_1 - p_2)^{n-x-y}}{\binom{n}{y} p_2^y (1 - p_2)^{n-y}}$$

$$= \binom{n-y}{x} \left(\frac{p_1}{1-p_2}\right)^x \left(\frac{1 - p_2 - p_1}{1 - p_2}\right)^{n-y-x}$$

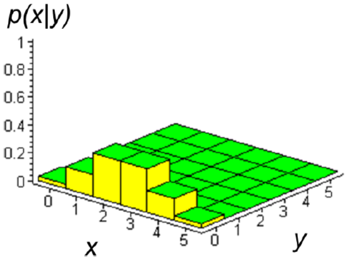
for integers $0 \leq x \leq n - y$

$\therefore X|Y = y \sim Bin\left(n - y, \frac{p_1}{1 - p_2}\right)$ for $y = 0, 1, \ldots, n$. By symmetry,

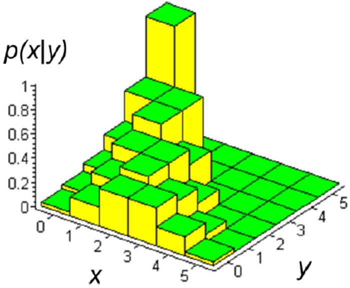
$Y|X = x \sim Bin\left(n - x, \frac{p_2}{1 - p_1}\right)$ for $x = 0, 1, \ldots, n$

For the trinomial distribution with parameters $n = 5, p_1 = p_2 = 1/3$:

$p_{X|Y}(x|0)$



All $p_{X|Y}(x|y)$ **for** $y = 0, 1, \ldots, 5$



There are totally 6 possible conditional distributions/pmf's of X given $Y = y$, with $y = 0, 1, \ldots, 5$, & all of them are different from one another

Conditional Distributions: Cont. Case I

Definition

Given the joint pdf $f_{X,Y}$ of 2 continuous r.v.'s X & Y, if $0 < f_Y(y) < \infty$ for some y, the *conditional prob density function of X given that $Y = y$* is defined by, for all values $x \in \mathbb{R}$.

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

or 0 otherwise.

- **The idea**: Only 1 r.v. X is random but Y is not random anymore because it has already happened to take on the value y & y is a constant throughout
- $f_{X,Y}(x,y)$ in the numerator is simply a function of x as y is known

Conditional Distributions: Cont. Case II

- $f_{X|Y}(x|y)$ is a proper pdf in x:
  - $f_{X|Y}(x|y) \geq 0$ for all x
  - $\int_{-\infty}^{\infty} f_{X|Y}(x|y)dx = 1$
- $X|Y = y$ is called a *conditional r.v.*
  - Just another cont. r.v. as those discussed before, which is defined from the reduced sample space consisting of only outcomes $\omega \in \Omega$, s.t. $Y(\omega) = y$ instead of the original sample space Ω
  - Take on values from the collection of all possible values of $X \Rightarrow$ Take on at most the same set of values as X
  - In general, different from X
  - Same as X only when X & Y are indept: $f_{X|Y}(x|y) = f_X(x)$ for all values $x, y \in \mathbb{R}$
- In differential notation: $f_{X|Y}(x|y)$ is defined as

$$P(x \leq X \leq x + dx | y \leq Y \leq y + dy) = \frac{f(x,y)dx\,dy}{f_Y(y)dy} = \frac{f(x,y)}{f_Y(y)}dx$$

Conditional Distributions: Cont. Case III

Multiplication Law & Law of Total Prob.

For any value $x, y \in \mathbb{R}$

- *Multiplication law*: The joint pdf can be re-expressed as

$$f_{X,Y}(x,y) = f_Y(y) f_{X|Y}(x|y) = f_X(x) f_{Y|X}(y|x)$$

- *Law of total prob.*: The marginal pdf's equal

$$f_X(x) = \int_{-\infty}^{\infty} f_Y(y) f_{X|Y}(x|y) dy \tag{5}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_X(x) f_{Y|X}(y|x) dx \tag{6}$$

Consider $f(x, y) = \frac{3}{2}x^2(1 - |y|)$, $-1 < x < 1$, $-1 < y < 1$, of which the marginal pdf's have been found in Example at Page 37. The conditional pdf's can be found as follows.

Solution: Given any $-1 < y < 1$, for $-1 < x < 1$,

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)} = \frac{\frac{3}{2}x^2(1 - |y|)}{1 - |y|} = \frac{3}{2}x^2 = f_X(x)$$

&, given any $-1 < x < 1$, for $-1 < y < 1$,

$$f_{Y|X}(y|x) = 1 - |y| = f_Y(y)$$

This justifies that X & Y are indept

Example: Conditional Dist: Cont. Case

Consider the *bivariate normal distribution* defined at Page 34: Assume that $\rho \neq 0$. With the marginal distribution obtained at Page 36, the *conditional pdf of X given $Y = y$* is, for $x, y \in \mathbb{R}$,

$$f_{X|Y}(x|y) = \frac{\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(z_x^2 + z_y^2 - 2\rho z_x z_y\right)\right]}{\frac{1}{\sqrt{2\pi}\sigma_Y} \exp\left[-\frac{1}{2}z_y^2\right]}$$

$$= \frac{1}{\sqrt{2\pi}\sigma_X\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(z_x^2 - 2\rho z_x z_y\right)\right]$$

$$\times \exp\left[-\frac{1}{2(1-\rho^2)}z_y^2 + \frac{1}{2}z_y^2\right]$$

Now, substituting back $z_x = (x - \mu_X)/\sigma_X$, & completing the square of the exponent in the first exponential term yield that $X|Y = y \sim N(\mu_X + \rho\sigma_X z_y, (1-\rho^2)\sigma_X^2)$

Conditional Distributions With ≥ 3 r.v.'s

When there are $m \geq 3$ r.v.'s of interest, say, $X_1, \cdots, X_m$, it is possible that one may wish to look at

- marginal distribution of any r.v. X_i
- joint distribution of a sub-vector of the n r.v.'s (e.g., joint distribution of $(X_1, X_2, \cdots, X_k)$, with $k < m$)
- conditional distribution of a random vector given an event concerning some other r.v.'s (e.g., distribution of $X_1, \cdots, X_k | X_{k+1} = x_{k+1}$, & distribution of $X_1, \cdots, X_k | X_{k+1} = x_{k+1}, \cdots, X_m = x_m$, with $k < m$)

Basically, one can always replace either X or Y or both X & Y by a random vector in all definitions & all the propositions & results regrading only 2 r.v.'s at a time discussed so far to address the above quantities of interest

Functions of Jointly Distributed r.v.'s I

- **Recall**: We discuss how to address the distribution of a function of 1 r.v. based on knowledge of the density of the r.v.
- Suppose that 2 jointly distributed cont. r.v.'s, X & Y, are transformed to another 2 cont r.v.'s, U & V, via the transformation

$$u = g_1(x, y)$$
$$v = g_2(x, y)$$

& that the transformation can be inverted to obtain the single-valued inverse

$$x = h_1(u, v)$$
$$y = h_2(u, v)$$

What is the joint pdf of U & V?

Change-of-Variable Technique

Let X & Y be jointly distributed cont. r.v.'s with pdf $f_{X,Y}$ with support $S \subset \mathbb{R}^2$. Let $U = g_1(X, Y)$ & $V = g_2(X, Y)$, where g_1 & g_2 have cont. partial derivatives, & there exist functions h_1 & h_2 s.t. $X = h_1(U, V)$ & $Y = h_2(U, V)$ for all values of X & Y. Then, *(U, V) has a joint pdf*

$$f_{U,V}(u, v) = \begin{cases} f_{X,Y}(h_1(u, v), h_2(u, v))|J^{-1}|, & (u, v) \in S^* \\ 0, & \text{otherwise} \end{cases}$$

where $|J^{-1}|$ is the absolute value of the reciprocal of

$$J = \text{determinant of} \begin{bmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial y} \\ \frac{\partial g_2}{\partial x} & \frac{\partial g_2}{\partial y} \end{bmatrix} = \frac{\partial g_1}{\partial x}\frac{\partial g_2}{\partial y} - \frac{\partial g_2}{\partial x}\frac{\partial g_1}{\partial y} \neq 0$$

for all x & y (J is called the *Jacobian*), &
$S^* = \{(u, v) \in \mathbb{R}^2 | u = g_1(x, y) \& v = g_2(x, y) \text{for some } (x, y) \in S\}$.

Consider the joint pdf $f_{X,Y}(x,y) = 2,\ 0 < x < y < 1$. What is the joint pdf of $U = X/Y$ & $V = Y$?

Solution: : It is easy to solve for x & y, namely, $x = uv$ & $y = v$, which result in a single–valued inverse of (x,y) from any (u,v). Computing $J = (y^{-1})(1) - (0)(-\ xy^{-2}) = y^{-1} = v^{-1}$ gives the joint pdf of U & V as

$$f_{U,V}(u,v) = f_{X,Y}(uv,v)|v| = 2v$$

for $0 < u, v < 1$
Remark: It is clear that X & Y are not indept as $X < Y$ in $f_{X,Y}$, but after the transformation, U & V are indept as $f_{U,V}$ can be factored into a product of functions involving only u & only v, respectively

Suppose that X & Y are indept uniform r.v.'s on $(0,1)$. Let

$$U = \sqrt{-2\ln X}\cos(2\pi Y) \qquad \& \qquad V = \sqrt{-2\ln X}\sin(2\pi Y)$$

Find the joint pdf of U & V.

Solution: Compute $J = \begin{vmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial y} \\ \frac{\partial g_2}{\partial x} & \frac{\partial g_2}{\partial y} \end{vmatrix} = \begin{vmatrix} \frac{u}{2x\ln x} & -2\pi v \\ \frac{v}{2x\ln x} & 2\pi u \end{vmatrix} = 2\pi e^{(u^2+v^2)/2}$

Since the joint pdf of X & Y is $f_{X,Y}(x,y) = f_X(x)f_Y(y) = 1,\ 0 < x, y < 1$, we have the joint pdf of U & V given by

$$f_{U,V}(u,v) = (1)\left(2\pi e^{(u^2+v^2)/2}\right)^{-1} = \frac{1}{\sqrt{2\pi}}e^{-u^2/2}\cdot\frac{1}{\sqrt{2\pi}}e^{-v^2/2}, \quad u, v \in \mathbb{R}$$

i.e., U & V are two indept $N(0,1)$ r.v.'s. This result is known as *Box–Muller transformation*, which says that we can *generate 2 indept draws/#s from a $N(0,1)$ r.v. by transforming 2 indept samples from a $U(0,1)$ r.v. by the above transformations*

Example: Functions of Jointly Distributed r.v.'s

Suppose that X & Y are indept gamma r.v.'s with parameters (α, λ) & (β, λ), respectively. Compute the joint pdf of $U = X + Y$ & $V = \dfrac{X}{X+Y}$

Solution:

1. U & V take values on $(0, \infty)$ & on $(0, 1)$, respectively
2. Unique solutions to equations $u = x + y$ & $v = x/(x + y)$ are $x = uv$ & $y = u(1 - v)$
3. $J = \begin{vmatrix} 1 & 1 \\ \frac{y}{(x+y)^2} & \frac{-x}{(x+y)^2} \end{vmatrix} = -\dfrac{1}{x+y} = -\dfrac{1}{u}$

$$f_{U,V}(u, v) = f_{X,Y}(uv, u(1 - v)) \times u$$
$$= \frac{\lambda e^{-\lambda u}(\lambda u)^{\alpha+\beta-1}}{\Gamma(\alpha+\beta)} \times \frac{v^{\alpha-1}(1-v)^{\beta-1}\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}, \quad u > 0, 0 < v < 1$$

$\therefore$ U & V are indep, where $U \sim G(\alpha + \beta, \lambda)$ & $V \sim B(\alpha, \beta)$

Sums of Jointly Distributed r.v.'s

One common usage of jointly distributed r.v.'s is about *sums of r.v.'s*

Sum of r.v.'s (Convolution)

- Discrete r.v.'s X & Y with joint pmf $p(x, y)$: The pmf of $Z = X + Y$ is

$$p_Z(z) = \sum_y p(z-y, y) = \sum_x p(x, z-x)$$

$$\overset{\text{indep of } X\&Y}{=} \sum_y p_X(z-y)p_Y(y) = \sum_x p_X(x)p_Y(z-x)$$

- Cont. r.v.'s $X\&Y$ with joint pmf $f(x, y)$: The pdf of $Z = X + Y$ is

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

$$\overset{\text{indep of } X\&Y}{=} \int_{-\infty}^{\infty} f_X(z-y f_Y(y)dy = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx$$

Show that *X + Y is a Poisson r.v.* with parameter $\lambda + \mu$ *if X & Y are indept Poisson r.v.'s* with parameters $\lambda > 0$ & $\mu > 0$, respectively

Solution: For $z = 0, 1, \ldots,$

$$p_Z(z) = \sum_x p_X(x) p_Y(z-x) = \sum_{x=0}^{z} \frac{e^{-\lambda}\lambda^x}{x!} \frac{e^{-\mu}\mu^{(z-x)}}{(z-x)!}$$

$$= e^{-(\lambda+\mu)}(\lambda+\mu)^z \sum_{x=0}^{z} \frac{1}{x!(z-x)!}\left(\frac{\lambda}{\lambda+\mu}\right)^x \left(\frac{\mu}{\lambda+\mu}\right)^{(z-x)}$$

$$= \frac{e^{-(\lambda+\mu)}(\lambda+\mu)^z}{z!}\left[\sum_{x=0}^{z} \frac{z!}{x!(z-x)!}\left(\frac{\lambda}{\lambda+\mu}\right)^x \left(1 - \frac{\lambda}{\lambda+\mu}\right)^{(z-x)}\right]$$

which reduces to the pmf of a $Poi(\lambda + \mu)$ r.v. as the latter sum (which sums up all probs of a $Bin(z, \lambda/(\lambda+\mu))$ r.v.) is 1

Example: Sum of 2 Independent Gamma r.v.'s

Assume that *X & Y are indept gamma r.v.'s* with parameters (α, λ) & (β, λ). Show that $Z = X + Y \sim G(\alpha+\beta, \lambda)$

Solution: For $z \leq 0$, $f_Z(z) = 0$, while, for $z > 0$,

$$
\begin{aligned}
f_Z(z) &= \int_0^z \left[\frac{\lambda^\alpha}{\Gamma(\alpha)}(z-y)^{\alpha-1}e^{-\lambda(z-y)}\right] \times \left[\frac{\lambda^\beta}{\Gamma(\beta)}y^{\beta-1}e^{-\lambda y}\right] dy \\
&= \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)}e^{-\lambda z}\int_0^z (z-y)^{\alpha-1}y^{\beta-1}\, dy \\
&= \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)}e^{-\lambda z}z^{\alpha+\beta-1}\int_0^1 (1-u)^{\alpha-1}u^{\beta-1}\, du \quad (\text{letting } u = y/z) \\
&= \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)}z^{\alpha+\beta-1}e^{-\lambda z}\left[\frac{\Gamma(\beta)\Gamma(\alpha)}{\Gamma(\alpha+\beta)}\right] = \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha+\beta)}z^{\alpha+\beta-1}e^{-\lambda z}
\end{aligned}
$$

where the latter integral is related to integrating the pdf of a $B(\alpha,\beta)$ r.v.

Assume that X & Y are indept $N(0, \sigma^2)$ & $N(0, 1)$ r.v.'s. Show that $Z = X + Y \sim N(0,\, 1 + \sigma^2)$

Solution: Consider the integrand in $f_Z(z)$, for $-\infty < z, y < \infty$,

$$\begin{aligned}
f_X(z-y)f_Y(y) &= \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(z-y)^2}{2\sigma^2}\right\}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{y^2}{2}\right) \\
&= \frac{1}{2\pi\sigma}\exp\left(-\frac{z^2}{2\sigma^2}\right)\exp\left\{-\frac{1+\sigma^2}{2\sigma^2}\left(y^2 - 2y\frac{z}{1+\sigma^2}\right)\right\} \\
&= \frac{1}{2\pi\sigma}\exp\left(-\frac{z^2}{2\sigma^2}\right)\exp\left\{\frac{z^2}{2\sigma^2(1+\sigma^2)}\right\} \\
&\quad \times \exp\left\{-\frac{1+\sigma^2}{2\sigma^2}\left(y - \frac{z}{1+\sigma^2}\right)^2\right\}
\end{aligned}$$

where the last equality follows from completing the square

Hence, letting $w = y - z/(1+\sigma^2)$,

$$
\begin{aligned}
f_Z(z) &= \frac{1}{2\pi\sigma} \exp\left\{-\frac{z^2}{2(1+\sigma^2)}\right\} \int_{-\infty}^{\infty} \exp\left(-\frac{1+\sigma^2}{2\sigma^2} w^2\right) dw \\
&= \frac{1}{2\pi\sigma} \exp\left\{-\frac{z^2}{2(1+\sigma^2)}\right\} \sqrt{\frac{2\pi\sigma^2}{1+\sigma^2}} \\
&\quad \times \int_{-\infty}^{\infty} \sqrt{\frac{1+\sigma^2}{2\pi\sigma^2}} \exp\left(-\frac{1+\sigma^2}{2\sigma^2} w^2\right) dw \\
&= \frac{1}{\sqrt{2\pi(1+\sigma^2)}} \exp\left\{-\frac{z^2}{2(1+\sigma^2)}\right\}, \qquad z \in \mathbb{R}
\end{aligned}
$$

which is the pdf of a $N(0, 1+\sigma^2)$ r.v.

Sometimes, we concern a collection of indep. cont. r.v.'s, especially on the maximum/minimum of them.

- The highest temperature in a month
- The smallest weight for the students in our class
- The maximal/minimal revenue for one company in one day
- $\cdots$

Definition

For n indept. r.v.'s $X_1, X_2, \cdots, X_n$, if we order them by

$$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)},$$

then $X_{(k)}$ is called the *kth-order statistic*.

pdf of Order Statistics

For n independently and identically distributed r.v.'s $X_1, X_2, \cdots, X_n$ with pdf f and cdf F, the density of $X_{(k)}$, the *kth-order statistic*, is

$$f_k(x) = \frac{n!}{(n-k)!(k-1)!} f(x) F^{k-1}(x)[1 - F(x)]^{n-k}.$$

Example: Order Statistics

- Minimum: $X_{(1)} = \min\{X_1, X_2, \cdots, X_n\}$

$$f_1(x) = nf(x)[1 - F(x)]^{n-1}.$$

- Maximum: $X_{(n)} = \max\{X_1, X_2, \cdots, X_n\}$

$$f_n(x) = nf(x)F^{n-1}(x).$$

- If n is odd, then $X_{(n+1)/2}$ is called the median of the X_i

$$f_{(n+1)/2}(x) = \frac{n!}{((n-1)/2)!((n-1)/2)!} f(x)F^{(n-1)/2}(x)[1-F(x)]^{(n-1)/2}.$$