Chapter 3: Probability calculation methods

Probability calculation methods

We now study two specific calculation methods relevant for probability and distribution calculations:

- transformations;
- sums of random variables.

Transformations

For a given random variable Y, we are interested in calculating the probability distribution of the *transformed* random variable

$$U = h(Y)$$

where $h(.)$ is a real-valued function with domain that includes the region $\mathcal{Y}$ on which Y takes values with positive probability.

Transformations (cont.)

Discrete case: In the discrete case, we consider how the probability attached to individual values is transported under the transformation.

Suppose Y takes values on the (countable) set

$$\{y_1, y_2, \ldots, y_n, \ldots\}.$$

Under the transformation $h(.)$, suppose the individual values map to the (countable) set

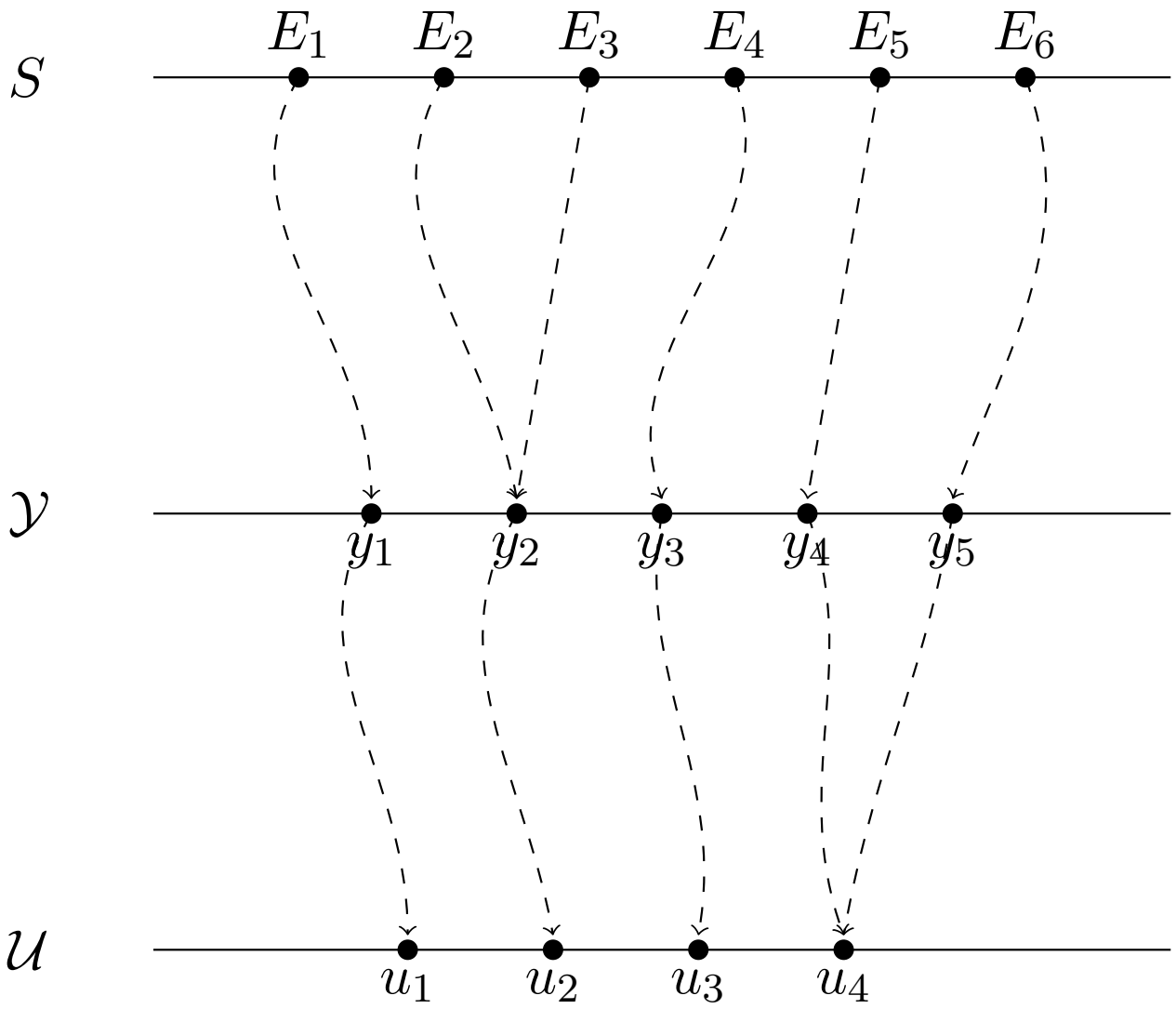
$$\{u_1, u_2, \ldots, u_n, \ldots\}$$

where

$$u_i = h(y_j)$$

for at least one j.

Transformations (cont.)



Transformations (cont.)

In the above example

$$u_1 = h(y_1) \quad u_2 = h(y_2) \quad u_3 = h(y_3) \quad u_4 = h(y_4) \text{ and } h(y_5)$$

We typically deal with simple transformations, for example

$$U = 2Y \qquad U = \exp(-Y) \qquad U = \frac{Y}{Y+1} \qquad U = Y^2$$

Transformations (cont.)

To derive the probability distribution of U, we simply see how the probability accumulates on individual values under the transformation.

In the picture above, we have four values $\{u_1, u_2, u_3, u_4\}$ which receive probability as follows:

$$p_U(u) = P(U = u) = \begin{cases} p_Y(y_1) & u = u_1 \\ p_Y(y_2) & u = u_2 \\ p_Y(y_3) & u = u_3 \\ p_Y(y_4) + p_Y(y_5) & u = u_4 \end{cases}$$

Transformations (cont.)

For 1-1 transformations, this is extremely straightforward to apply; we simply

- identify the set of values $\{u_1, u_2, \ldots, u_n, \ldots\}$ that the values $\{y_1, y_2, \ldots, y_n, \ldots\}$ map onto – in this case there is a one-to-one correspondence between the sets, and

$$u_i = h(y_i);$$

- for each i, assign the probability $p_Y(y_i)$ to u_i to obtain $p_U(u_i)$

Transformations (cont.)

Example (Poisson case)

Suppose that Y has a Poisson distribution with parameter λ. Let $U = 3Y$; then only the values in the set

$$\mathcal{U} = \{0, 3, 6, 9, \ldots, 3n, \ldots\}$$

receive any probability, and

$$p_U(u) = \frac{\lambda^{u/3}}{(u/3)!} e^{-\lambda} \qquad u \in \{0, 3, 6, 9, \ldots\}$$

with $p_U(u)$ equal to zero for all other values of u.

Transformations (cont.)

If the transformation is not 1-1, we essentially have to identify the set $\mathcal{U}$, and compute for each case the probability assigned to the values in the set.

Transformations (cont.)

Continuous case: In the continuous random variable case, we try to adopt the same strategy and follow how values in the set $\mathcal{Y}$ carry their probability to values in a set $\mathcal{U}$.

The obstacle is that for a continuous random variable Y we have

$$P(Y = y) = 0$$

for all real values y.

Transformations (cont.)

To get around this, we can instead consider

$$P(Y \in \mathcal{A})$$

for a subset (eg interval) of the real line. For example, we could consider $\mathcal{A} = [2, 5]$, and the probability

$$P(Y \in \mathcal{A}) \equiv P(2 \leq Y \leq 5).$$

Transformations (cont.)

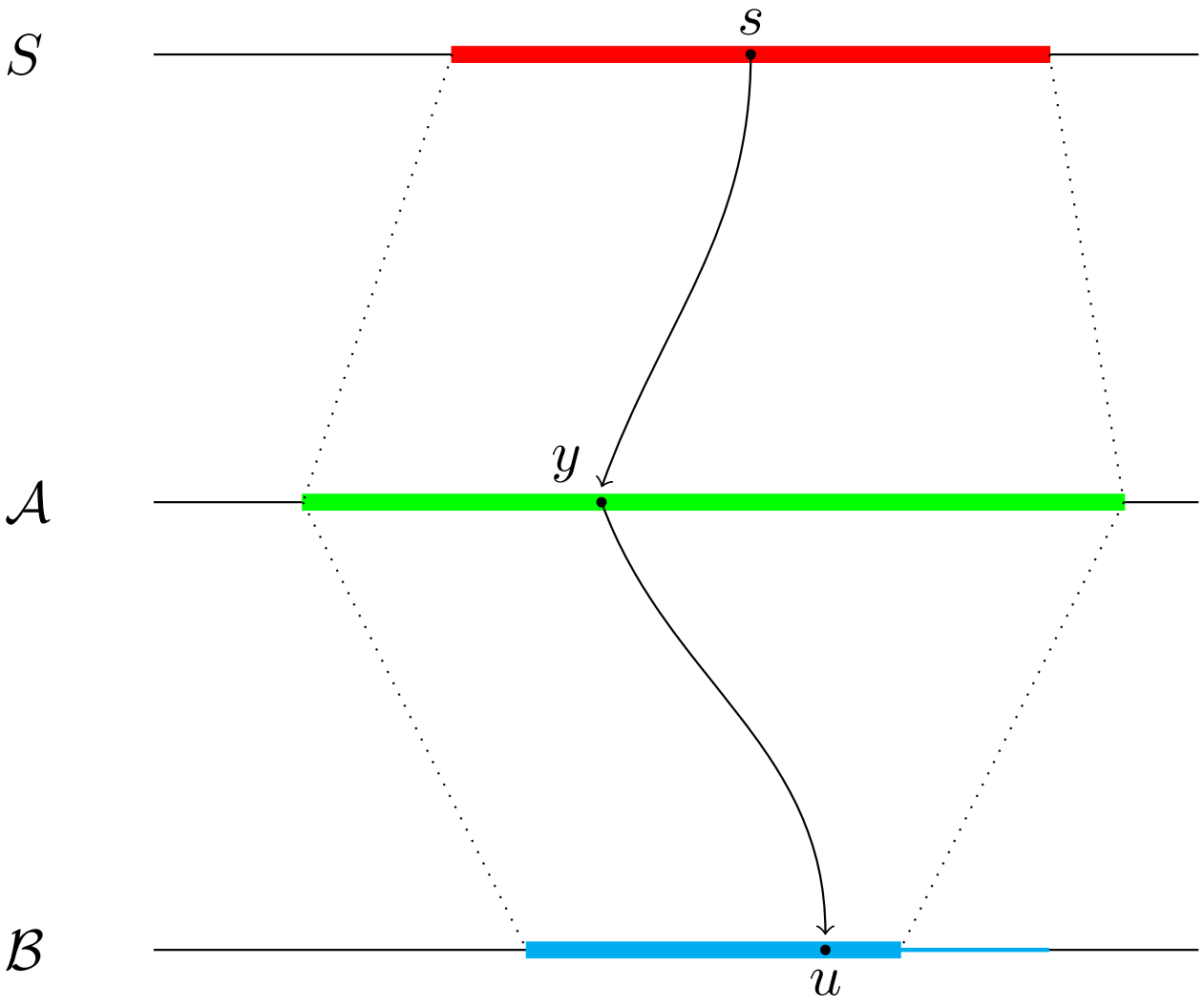
Suppose that $\mathcal{B}$ represents the set of values containing all the points that points in $\mathcal{A}$ can be mapped to under the transformation $h(.)$. That is,

$$\mathcal{B} = \{u : u = h(y), \text{some } y \in \mathcal{A}\}$$

Then we have that probability is preserved under the mapping

$$P(Y \in \mathcal{A}) = P(U \in \mathcal{B}).$$

Transformations (cont.)



Transformations (cont.)

Example

Suppose Y is a continuous random variable that can take values on $[0, \infty)$, and $U = h(Y) = 2Y$. This means that we need to consider the set

$$\mathcal{U} = \{u : u = 2y, \text{ some } y \in [0, \infty)\} \equiv [0, \infty)$$

that is, every point y gets mapped to $u = 2y$.

Transformations (cont.)

Example

Consider interval $\mathcal{A} = [a, b]$ for $a < b$; if we let $c = h(a) = 2a$ and $d = 2b$, then we have that

$$P(a \leq Y \leq b) = P(c \leq U \leq d).$$

In terms of the cdfs, we have that

$$F_Y(b) - F_Y(a) = F_U(d) - F_U(c)$$

and letting $a \longrightarrow 0$, we have that

$$F_Y(b) = P(Y \leq b) = P(U \leq d) = F_U(d).$$

Transformations (cont.)

Example

That is, for $b \in \mathbb{R}$.

$$F_Y(b) = F_U(2b)$$

Now, given F_Y, we wish to compute F_U. We can do this by noting that, for arbitrary u

$$F_U(u) = P(U \leq u) = P(U/2 \leq u/2) = P(Y \leq u/2) = F_Y(u/2).$$

Transformations (cont.)

Example

To get to the pdf, we need to differentiate both sides with respect to u:

$$f_U(u) = \frac{dF_U(u)}{du} = \frac{dF_Y(u/2)}{du} = \frac{1}{2} f_Y(u/2)$$

by the chain rule. For example, if $F_Y(y) = 1 - \exp\{-y\}$, then for $u \geq 0$,

$$f_U(u) = \frac{d}{du}\left\{1 - \exp\{-u/2\}\right\} = \frac{1}{2}\exp\{-u/2\}.$$

Transformations (cont.)

The key relationship here is that

$$P(Y \in \mathcal{A}) = P(U \in \mathcal{B}).$$

We want to compute $F_U(.)$ or $f_U(.)$, which means we want to fix $\mathcal{B}$ first. Therefore, write

$$h^{-1}(\mathcal{B})$$

for the set of y values that map onto $\mathcal{B}$. Then we have

$$P(U \in \mathcal{B}) = P(Y \in h^{-1}(\mathcal{B}))$$

Transformations (cont.)

If $h(.)$ is a 1–1 transformation, then $h^{-1}(.)$ is the inverse transformation which we can often compute explicitly.

$h(y)$	$h^{-1}(y)$
$y + \theta$	$y - \theta$
λy	y/λ
e^y	$\ln y$
$\dfrac{y}{1+y}$	$\dfrac{y}{1-y}$

etc.

Transformations (cont.)

For 1–1 (or monotonic) transformations, we consider two cases

1. $h(y)$ increasing: eg $U = 2X$
2. $h(y)$ decreasing: eg $U = 2/X$

Transformations (cont.)

For a 1–1 increasing transformation, we have for the cdf

$$F_U(u) = P(U \leq u) = P(h(Y) \leq u) = P(Y \leq h^{-1}(u))$$
$$= F_Y(h^{-1}(u))$$

and therefore for the pdf, by the chain rule

$$f_U(u) = \frac{dF_U(u)}{du} = \frac{dF_Y(h^{-1}(u))}{du} = f_Y(h^{-1}(u))\frac{dh^{-1}(u)}{du}.$$

Transformations (cont.)

For a 1–1 decreasing transformation, we have instead

$$F_U(u) = P(U \leq u) = P(h(Y) \leq u) = P(Y \geq h^{-1}(u))$$
$$= 1 - F_Y(h^{-1}(u))$$

and therefore for the pdf, by the chain rule

$$f_U(u) = \frac{dF_U(u)}{du} = -\frac{dF_Y(h^{-1}(u))}{du} = -f_Y(h^{-1}(u))\frac{dh^{-1}(u)}{du}.$$

Transformations (cont.)

Therefore we can write

$$f_U(u) = f_Y(h^{-1}(u)) \left| \frac{dh^{-1}(u)}{du} \right|$$

to cover both cases. The term

$$\left| \frac{dh^{-1}(u)}{du} \right|$$

is the *Jacobian* of the transformation.

Transformations (cont.)

Note

The identity

$$P(U \in \mathcal{B}) = P(Y \in h^{-1}(\mathcal{B}))$$

can be re-written

$$\int_{\mathcal{B}} f_U(u)\, du = \int_{h^{-1}(\mathcal{B})} f_Y(y)\, dy$$

Transformations (cont.)

Note (cont.)

Changing variables in the right-hand integral

$$u = h(y) \quad \Longleftrightarrow \quad y = h^{-1}(u)$$

we have by the rules of integration

$$\int_{\mathcal{B}} f_U(u)\, du = \int_{\mathcal{B}} f_Y(h^{-1}(u)) \left| \frac{dh^{-1}(u)}{du} \right| du$$

and the same result follows by equating the integrands.

Transformations (cont.)

If $h(.)$ is not 1–1 (non-monotonic), then the identity

$$P(U \in \mathcal{B}) = P(Y \in h^{-1}(\mathcal{B}))$$

still holds, we just need more care in computing the right-hand side.

Transformations (cont.)

Example

Suppose that Y has a continuous uniform distribution on the range $[0, 1]$, and that

$$U = Y(1 - Y)$$

that is $U = h(Y)$ with $h(t) = t(1 - t)$.

First, as $\mathcal{Y} = [0, 1]$, we have that

$$\mathcal{U} = [0, 1/4].$$

Transformations (cont.)

Example

Then for $0 \leq u \leq 1/4$, we have that

$$P(U \leq u) = P(Y(1-Y) \leq u) = P(Y - Y^2 - u \leq 0)$$

That is

$$F_U(u) = P(Y^2 - Y + u \geq 0) = P((Y - y_1)(Y - y_2) \geq 0)$$

where, factorizing the quadratic, we have

$$(y_1, y_2) = \left(\frac{1 - \sqrt{1-4u}}{2}, \frac{1 + \sqrt{1-4u}}{2}\right)$$

Transformations (cont.)

Example

Now, in the quadratic

$$(y - y_1)(y - y_2) \geq 0 \qquad \iff y \leq y_1 \quad \text{or} \quad y \geq y_2$$

so for $0 \leq u \leq 1/4$

$$\begin{aligned} F_U(u) &= P((Y - y_1)(Y - y_2) \geq 0) \\ &= P(Y \leq y_1) + P(Y \geq y_2) \\ &= F_Y(y_1) + (1 - F_Y(y_2)) \\ &= y_1 + 1 - y_2 \\ &= 1 - \sqrt{1 - 4u}. \end{aligned}$$

Transformations (cont.)

Example

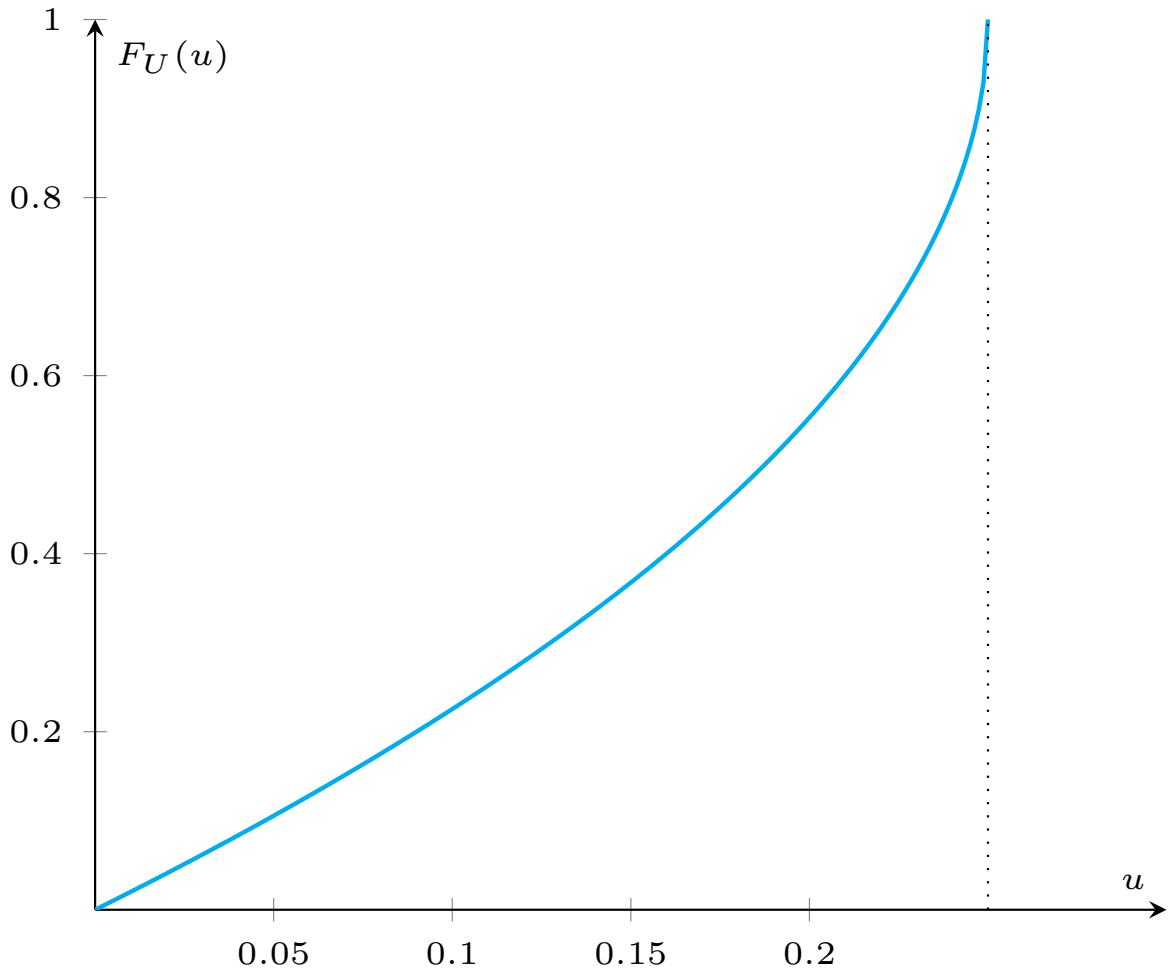
Differentiating, we have

$$f_U(u) = \frac{2}{\sqrt{1-4u}} \qquad 0 \le u \le 1/4.$$

For $u < 0$ and $u > 1/4$, $f_U(u) = 0$.

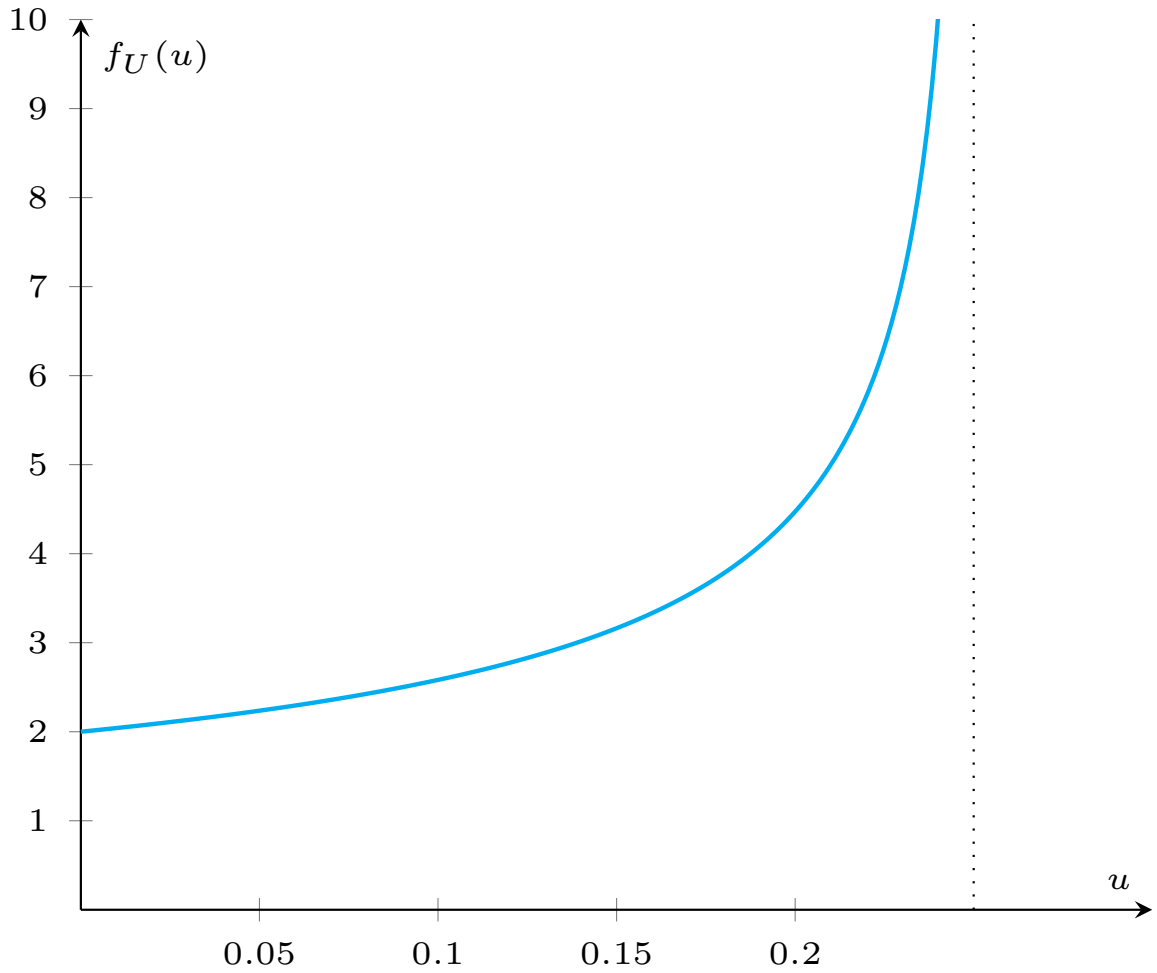
Transformations (cont.)

$$F_U(u) = 1 - \sqrt{1 - 4u}.$$



Transformations (cont.)

$$f_U(u) = 2/\sqrt{1-4u}.$$



Transformations (cont.)

Note

For non 1-1 transformations, it is usually necessary to go back to 'first principles' and investigate

$$P(U \in \mathcal{B}) = P(Y \in h^{-1}(\mathcal{B}))$$

for $\mathcal{B} = (-\infty, u]$ and then carefully construct $h^{-1}(\mathcal{B})$ from it.

Transformations (cont.)

Expectations

We previously saw for continuous random variable Y and function $g(.)$

$$\mathbb{E}[g(Y)] = \int_{-\infty}^{\infty} g(y) f_Y(y)\, dy.$$

However, $U = g(Y)$ is a random variable in its own right. We have

$$\mathbb{E}[U] = \int_{-\infty}^{\infty} u f_Y(U)\, du.$$

From the above argument, we see that

$$\mathbb{E}[U] = \mathbb{E}[g(Y)]$$

that is, we can regard the computation in two equivalent ways.

Transformations (cont.)

Moment Generating Functions

Recall the definition of the moment generating function (mgf): for $-b \leq t \leq b$, $b > 0$,

$$m_Y(t) = \mathbb{E}[e^{tY}].$$

If $U = cY + d$, for constants $c > 0$ and d, we have that

$$m_U(t) = \mathbb{E}\left[e^{tU}\right] = \mathbb{E}\left[e^{t(cY+d)}\right] = \mathbb{E}\left[e^{(ct)Y}e^{dt}\right] = e^{dt}\mathbb{E}\left[e^{(ct)Y}\right]$$

as e^{dt} is a constant. Therefore

$$m_U(t) = e^{dt}m_Y(ct).$$

Transformations (cont.)

Moment Generating Functions: Normal case

If Y has a standard Normal distribution,

$$m_Y(t) = \exp\left\{\frac{t^2}{2}\right\} \qquad t \in \mathbb{R}.$$

If $U = \sigma Y + \mu$, then

$$m_U(t) = e^{\mu t} m_Y(\sigma t) = \exp\left\{\mu t + \frac{(\sigma t)^2}{2}\right\}.$$

Thus, by the earlier result for the transformations of standard Normal random variables

$$m(t) = \exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\}$$

is the mgf of the $Normal(\mu, \sigma^2)$ distribution.

Sums of random variables

Example (Two dice)

Two fair dice are rolled, with the outcomes of the rolls independent events.

What is the probability that the sum of the scores of the two dice exceeds 10 ?

Sums of random variables (cont.)

Example (Two dice)

Let Y_1 and Y_2 be the random variables recording the scores on the two dice: Y_1 and Y_2 have discrete uniform distributions on $\{1, 2, 3, 4, 5, 6\}$.

	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

The outcome combinations in this table are equally likely, and occur with probability 1/36.

Sums of random variables (cont.)

Example (Two dice)

Let Y denote the total score. The possible total scores are

$$\mathcal{Y} = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

and to compute $P(Y = y)$ we can sum down the antidiagonals of the table to obtain

$$p(y) = P(Y = y) = \frac{\min\{6, y-1\} - \max\{1, y-6\} + 1}{36} \quad y \in \mathcal{Y}$$

Sums of random variables (cont.)

Example (Two dice)

That is

$$p(y) = \begin{cases} 1/36 & y = 2, 12 \\ 2/36 & y = 3, 11 \\ 3/36 & y = 4, 10 \\ 4/36 & y = 5, 9 \\ 5/36 & y = 6, 8 \\ 6/36 & y = 7. \end{cases}$$

Therefore

$$P[Y > 10] = P[Y = 11] + P[Y = 12] = \frac{3}{36} = \frac{1}{12}.$$

Sums of random variables (cont.)

In the example, we are computing the probability distribution of

$$Y = Y_1 + Y_2$$

under the assumption that Y_1 and Y_2 are *independent random variables*, that is,

- they are connected to events that are independent in the original sample space.

Sums of random variables (cont.)

The assumption of independence of the events means that we must have

$$P((Y_1 = y_1) \cap (Y_1 = y_2)) = P(Y_1 = y_1) \times P(Y_2 = y_2)$$

for all pairs of possible realizations (y_1, y_2), each value taken from the set $\{1, 2, 3, 4, 5, 6\}$.

Sums of random variables (cont.)

Now consider the event $(Y = y)$ for the set of possible values

$$y \in \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\};$$

we have the partition given by result of the first roll

$$(Y = y) \equiv \bigcup_{y_1=1}^{6} (Y = y \cap Y_1 = y_1).$$

– that is, when we observe $Y = 5$, say, that could arise as the unordered pairs

$$(1, 4), (2, 3), (3, 2), (4, 1).$$

Sums of random variables (cont.)

However the event

$$(Y = y \cap Y_1 = y_1)$$

can be re-written

$$(Y_2 = y - y_1 \cap Y_1 = y_1)$$

as if $Y = y$ and $Y_1 = y_1$, we **must** have $Y_2 = y - y_2$.

Sums of random variables (cont.)

Using Axiom III, we can then write

$$
\begin{aligned}
P(Y = y) &= \sum_{y_1=1}^{6} P((Y = y) \cap (Y_1 = y_1)) \\
&= \sum_{y_1=1}^{6} P((Y_2 = y - y_2) \cap (Y_1 = y_1)) \\
&= \sum_{y_1=1}^{6} P(Y_1 = y_1) P(Y_2 = y - y_1)
\end{aligned}
$$

Sums of random variables (cont.)

For the range of the summation, we require for fixed y that

$$1 \leq y_1 \leq 6$$

so that $P(Y_1 = y_1) > 0$ and

$$1 \leq y - y_1 \leq 6$$

in order to make

$$P(Y_2 = y - y_1) > 0.$$

Sums of random variables (cont.)

That is we need to have

$$1 \leq y_1 \leq 6$$

$$y - 6 \leq y_1 \leq y - 1$$

in order for a term in the sum to be non-zero.

Sums of random variables (cont.)

Therefore we need

$$\max\{1, y-6\} \leq y_1 \leq \min\{6, y-1\}$$

That is for example,

$$\begin{aligned} y = 3: &\qquad 1 \leq y_1 \leq 2 \\ y = 7: &\qquad 1 \leq y_1 \leq 6 \\ y = 10: &\qquad 4 \leq y_1 \leq 6 \end{aligned}$$

and so on.

Sums of random variables (cont.)

Therefore

$$
\begin{aligned}
P(Y = y) &= \sum_{y_1 = \max\{1, y-6\}}^{\min\{6, y-1\}} P(Y_1 = y_1) P(Y_2 = y - y_1) \\
&= \sum_{y_1 = \max\{1, y-6\}}^{\min\{6, y-1\}} \frac{1}{6} \times \frac{1}{6} \\
&= \frac{\min\{6, y-1\} - \max\{1, y-6\} + 1}{36}
\end{aligned}
$$

Sums of random variables (cont.)

The crucial components of this calculation are

- the partition of the event $(Y = y)$ according to the possible values of Y_1;
- the theorem of total probability;
- the independence of the outcomes of the two rolls, equivalently, the independence of the random variables Y_1 and Y_2.

Using these elements, we can derive a general result in the discrete case.

Sums of random variables (cont.)

Suppose Y_1 and Y_2 are discrete random variables, with Y_1 and Y_2 independent. Suppose that Y_1 and Y_2 take values on the sets $\mathcal{Y}_1$ and $\mathcal{Y}_2$ respectively. Let $Y = Y_1 + Y_2$. Then

$$\begin{aligned}
P(Y = y) &= P\left(\bigcup_{y_1 \in \mathcal{Y}_1} (Y = y) \cap (Y_1 = y_1)\right) \\
&= \sum_{y_1 \in \mathcal{Y}_1} P((Y = y) \cap (Y_1 = y_1)) \\
&= \sum_{y_1 \in \mathcal{Y}_1} P((Y_2 = y - y_1) \cap (Y_1 = y_1)) \\
&= \sum_{y_1 \in \mathcal{Y}_1} P(Y_2 = y - y_1) P(Y_1 = y_1)
\end{aligned}$$

Sums of random variables (cont.)

That is

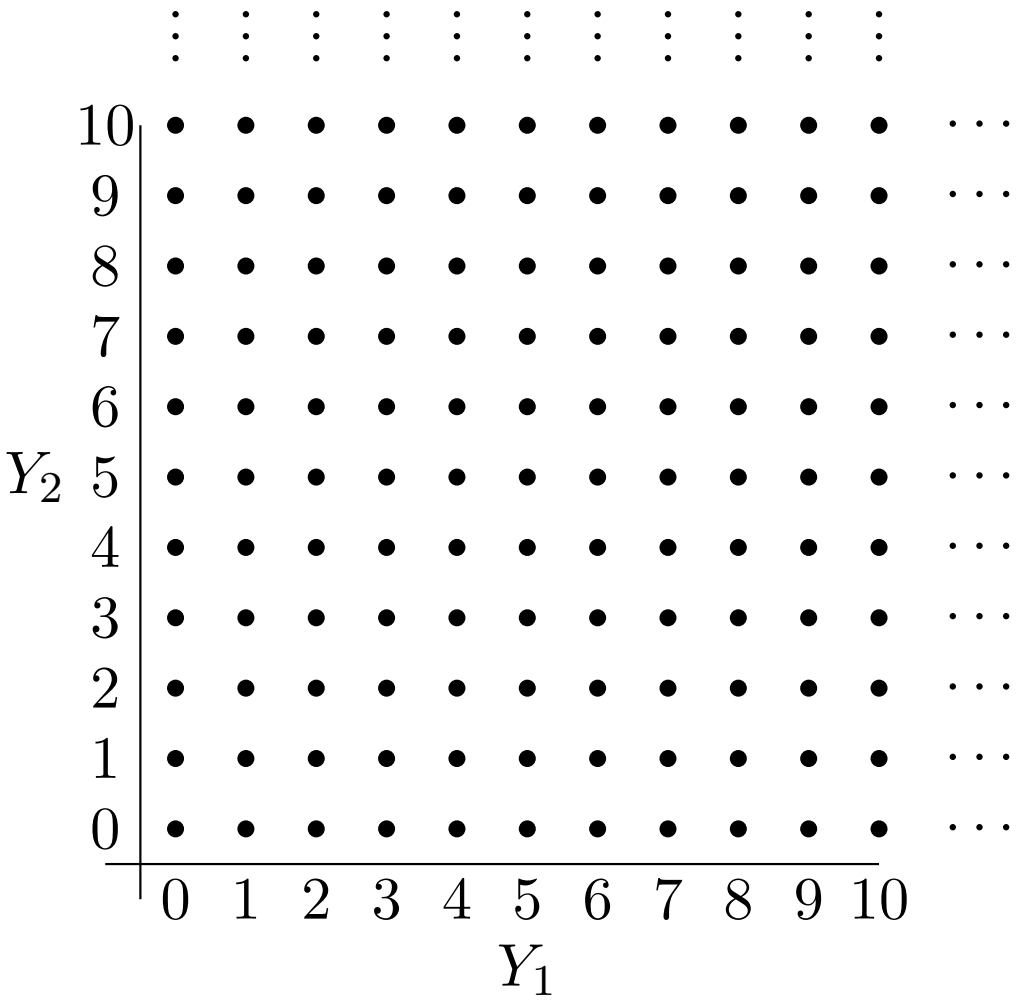
$$P(Y = y) = \sum_{y_1 \in \mathcal{Y}_1^*} P(Y_2 = y - y_1) P(Y_1 = y_1)$$

where

$$\mathcal{Y}_1^* = \{y_1 : \ y_1 \in \mathcal{Y}_1 \text{ and } y - y_1 \in \mathcal{Y}_2\}$$

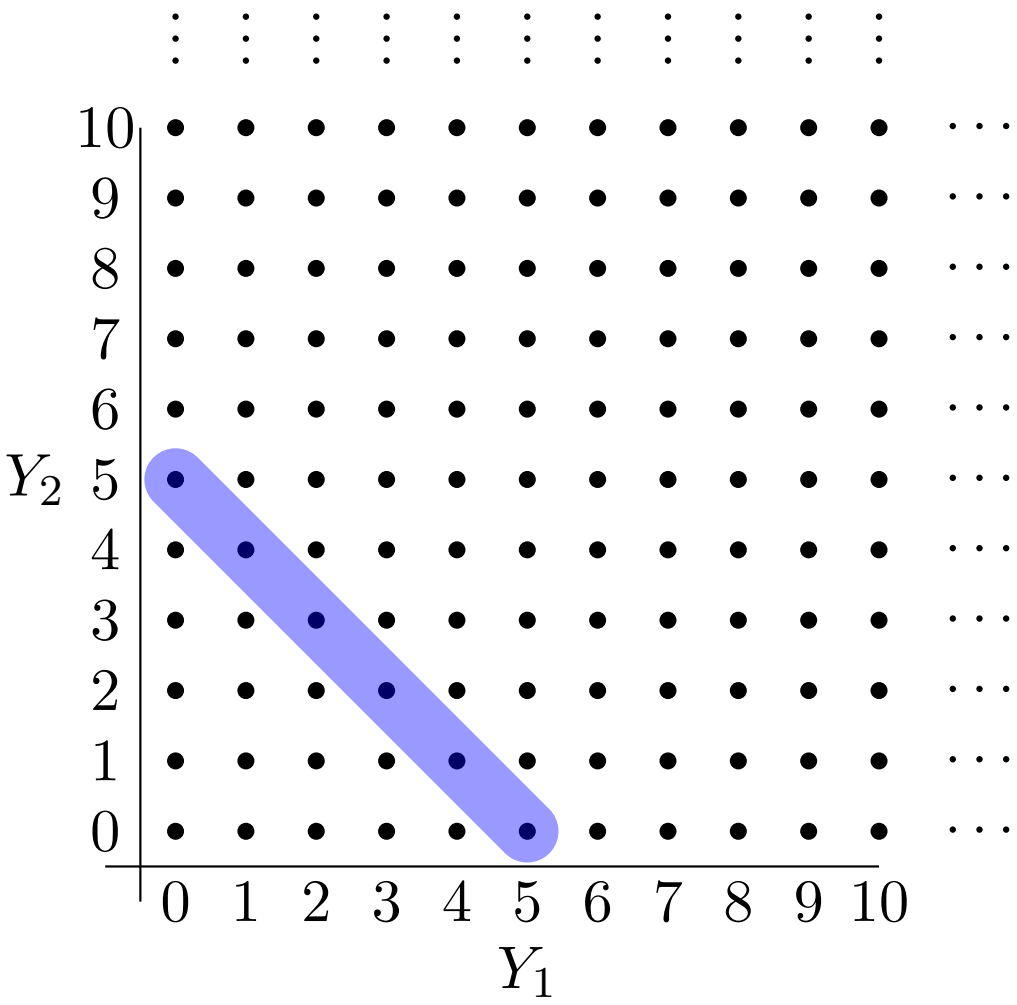
Sums of random variables (cont.)

Y_1, Y_2 take values on $\{0, 1, 2, \ldots\}$.



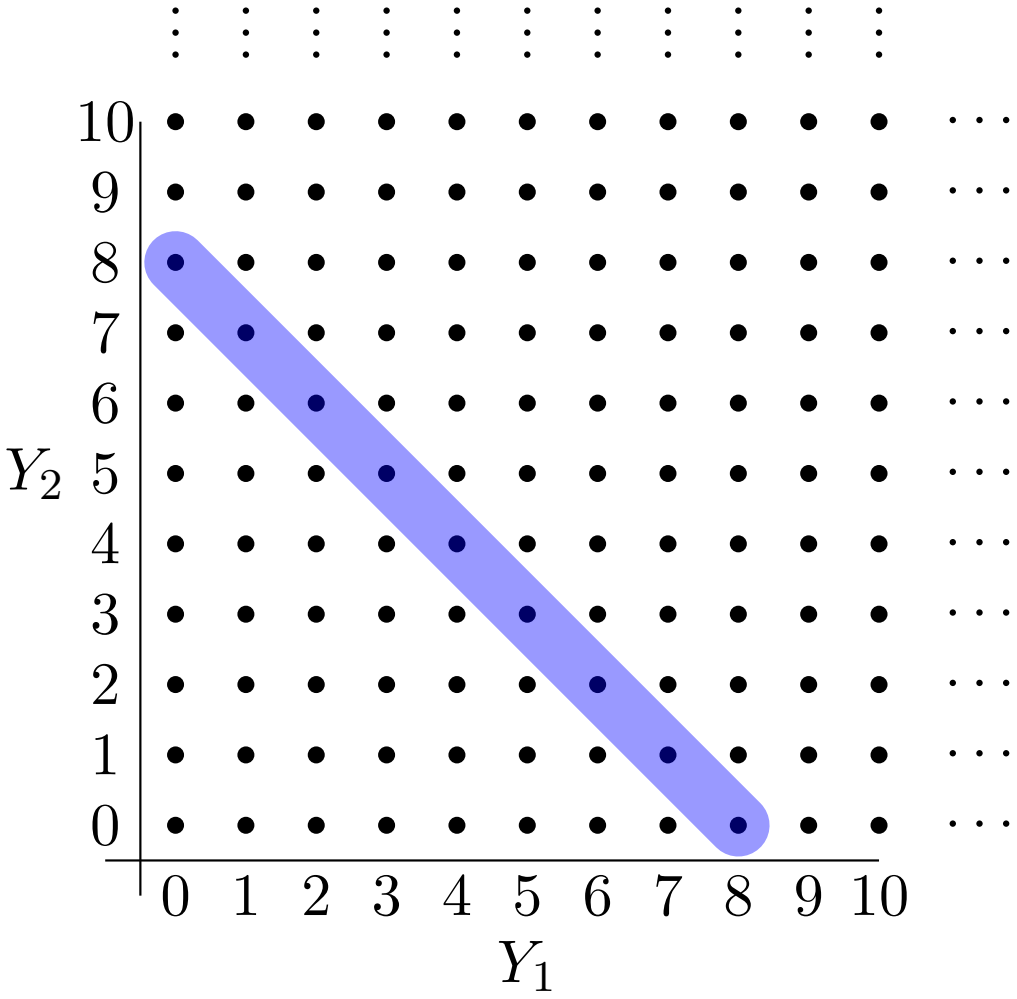
Sums of random variables (cont.)

$Y_1 + Y_2 = 5$



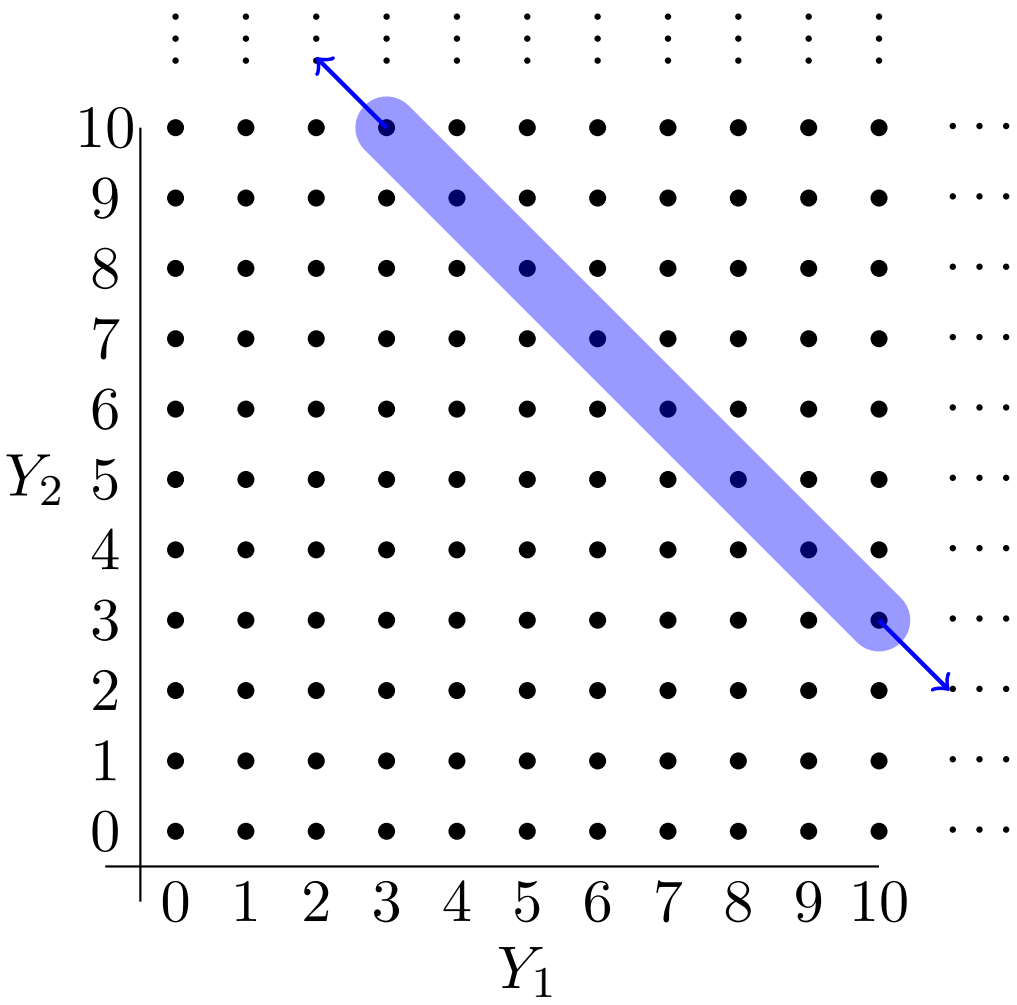
Sums of random variables (cont.)

$Y_1 + Y_2 = 8$



Sums of random variables (cont.)

$Y_1 + Y_2 = 13$



Sums of random variables (cont.)

In this example, for $y = 0, 1, 2, \ldots$

$$P(Y = y) = \sum_{y_1=0}^{y} P(Y_1 = y_1) P(Y_2 = y - y_1)$$

Sums of random variables (cont.)

If Y_1 and Y_2 are independent Poisson random variables with parameters λ_1 and λ_2, then if $Y = Y_1 + Y_2$,

$$
\begin{aligned}
P(Y = y) &= \sum_{y_1=0}^{y-y_1} P(Y_1 = y_1) P(Y_2 = y - y_2) \\
&= \sum_{y_1=0}^{y-y_1} \frac{\lambda_1^{y_1} e^{-\lambda_1}}{y_1!} \frac{\lambda_2^{y-y_1} e^{-\lambda_2}}{(y-y_1)!} \\
&= e^{-(\lambda_1+\lambda_2)} \lambda_2^y \sum_{y_1=0}^{y-y_1} \left(\frac{\lambda_1}{\lambda_2}\right)^{y_1} \frac{1}{y_1!(y-y_1)!} \\
&= \frac{e^{-(\lambda_1+\lambda_2)} \lambda_2^y}{y!} \left(1 + \frac{\lambda_1}{\lambda_2}\right)^y
\end{aligned}
$$

Sums of random variables (cont.)

But

$$\lambda_2^y \left(1 + \frac{\lambda_1}{\lambda_2}\right)^y = (\lambda_1 + \lambda_2)^y$$

so therefore

$$P(Y = y) = \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1 + \lambda_2)^y}{y!} \qquad y = 0, 1, 2, \ldots$$

that is, Y has a Poisson distribution with parameter $\lambda_1 + \lambda_2$.

Sums of random variables (cont.)

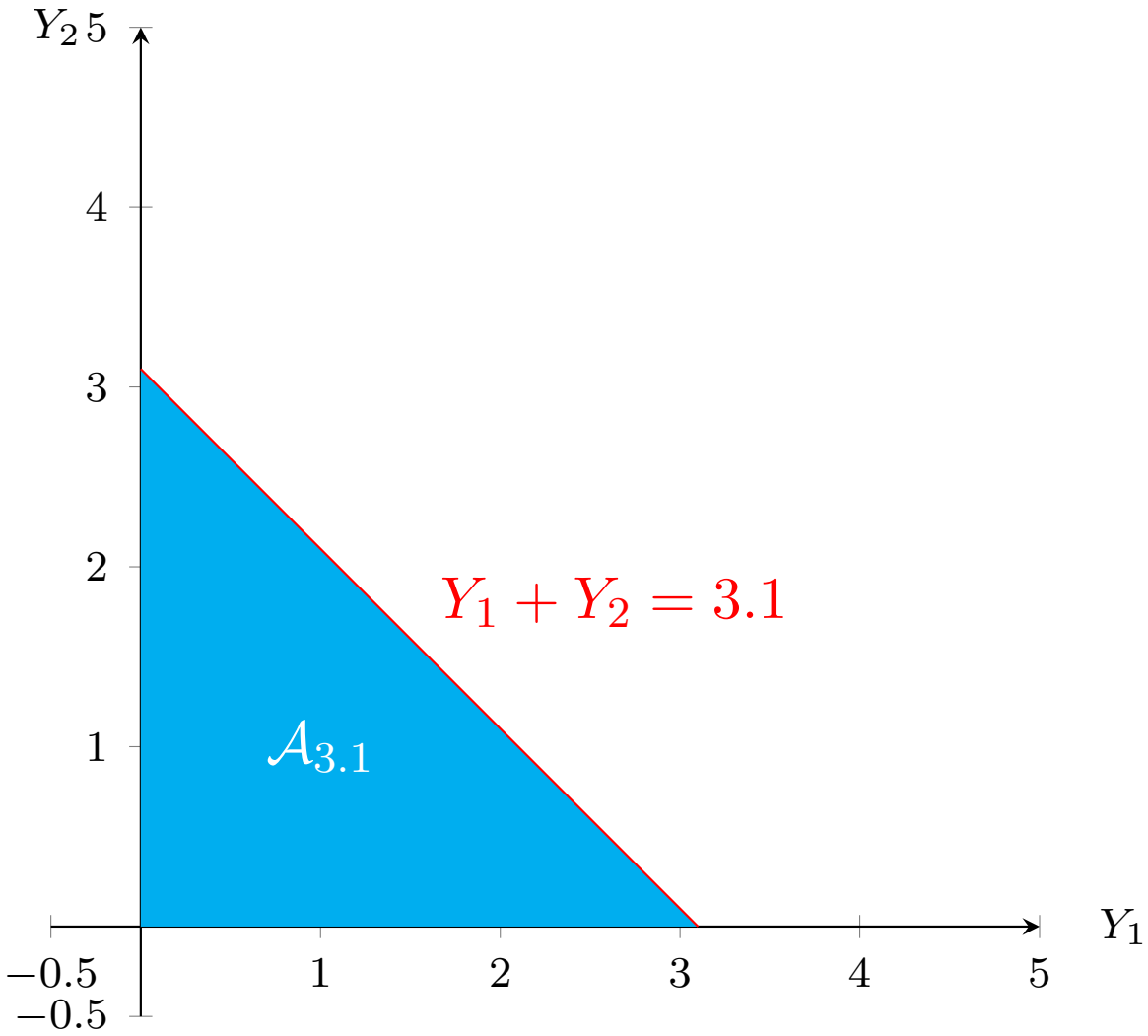
If Y_1 and Y_2 are independent continuous random variables, a similar strategy can be used. As before for continuous random variables, we start with the cdf and aim to compute

$$F_Y(y) = P(Y \leq y)$$

If Y_1 and Y_2 are continuous, then it follows that Y is continuous.

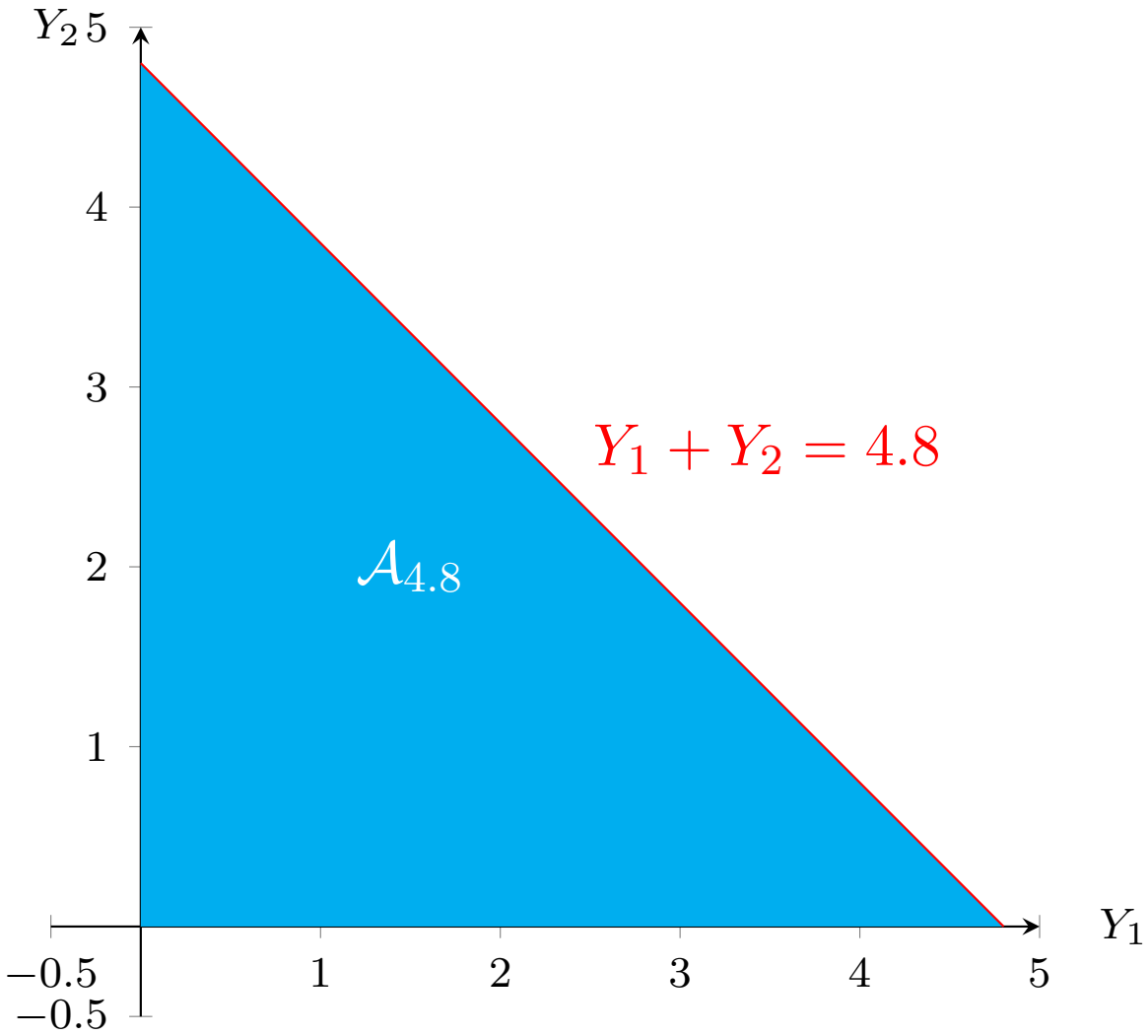
Sums of random variables (cont.)

Y_1, Y_2 taking values on $\mathbb{R}^+$; region $\mathcal{A}_{3.1}$ is where $Y_1 + Y_2 \leq 3.1$.



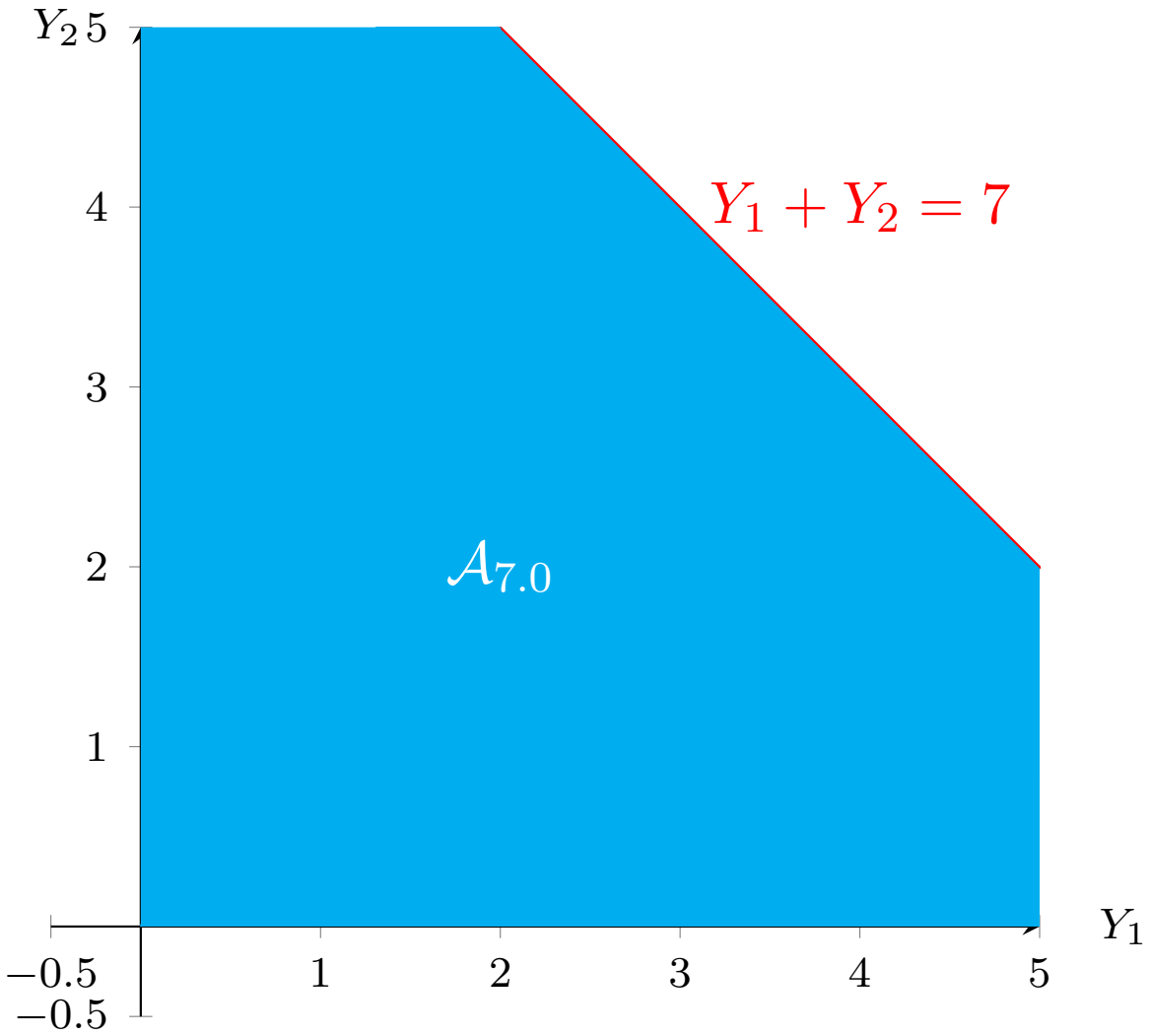
Sums of random variables (cont.)

Y_1, Y_2 taking values on $\mathbb{R}^+$; region $\mathcal{A}_{4.8}$ is where $Y_1 + Y_2 \leq 4.8$.



Sums of random variables (cont.)

Y_1, Y_2 taking values on $\mathbb{R}^+$; region $\mathcal{A}_{7.0}$ is where $Y_1 + Y_2 \leq 7$.



Sums of random variables (cont.)

Then we have by definition that

$$F_Y(y) = P(Y \leq y) = \int_{-\infty}^{y} f_Y(t)\, dt$$

where $f_Y(y)$ is the pdf for Y. However,

$$(Y \leq y) = (Y_1 + Y_2 \leq y)$$

Therefore to compute $F_Y(y) = P(Y \leq y)$ we need to aggregate all the probability across the blue shaded region, denoted $\mathcal{A}_y$, in the above pictures.

Sums of random variables (cont.)

In the discrete case in which

$$p_Y(y) = P(Y = y) = \sum_{y_1 \in \mathcal{Y}_1^*} P(Y_2 = y - y_1) P(Y_1 = y_1)$$

$$= \sum_{y_1 \in \mathcal{Y}_1^*} p_{Y_2}(y - y_1) p_{Y_1}(y_1)$$

where

- $p_Y(.)$ is the pmf for Y,
- $p_{Y_1}(.)$ is the pmf for Y_1,
- $p_{Y_2}(.)$ is the pmf for Y_2.

Sums of random variables (cont.)

we may write

$$f_Y(y) = \int_{-\infty}^{\infty} f_{Y_2}(y - y_1) f_{Y_1}(y_1)\, dy_1$$

where

- $f_Y(.)$ is the pdf for Y,
- $f_{Y_1}(.)$ is the pdf for Y_1,
- $f_{Y_2}(.)$ is the pdf for Y_2.

Sums of random variables (cont.)

Note

This is termed a *convolution* formula.

Sums of random variables (cont.)

Example (Exponential random variables)

Suppose

$$Y_1 : f_{Y_1}(y_1) = \frac{1}{\beta_1} e^{-y_1/\beta_1} \qquad y_1 \geq 0$$

$$Y_2 : f_{Y_2}(y_2) = \frac{1}{\beta_2} e^{-y_2/\beta_2} \qquad y_2 \geq 0$$

are independent random variables, and let $Y = Y_1 + Y_2$. For convenience, write

$$\lambda_1 = \frac{1}{\beta_1} \qquad \lambda_2 = \frac{1}{\beta_2}.$$

Sums of random variables (cont.)

Example (Exponential random variables)

Then from the convolution formula, for $y \geq 0$,

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{Y_2}(y - y_1) f_{Y_1}(y_1)\, dy_1 \\ &= \int_0^y \lambda_2 e^{-\lambda_2(y-y_1)} \lambda_1 e^{-\lambda_1 y_1}\, dy_1 \\ &= \lambda_1 \lambda_2 e^{-\lambda_2 y} \int_0^y e^{-(\lambda_1 - \lambda_2) y_1}\, dy_1 \end{aligned}$$

Sums of random variables (cont.)

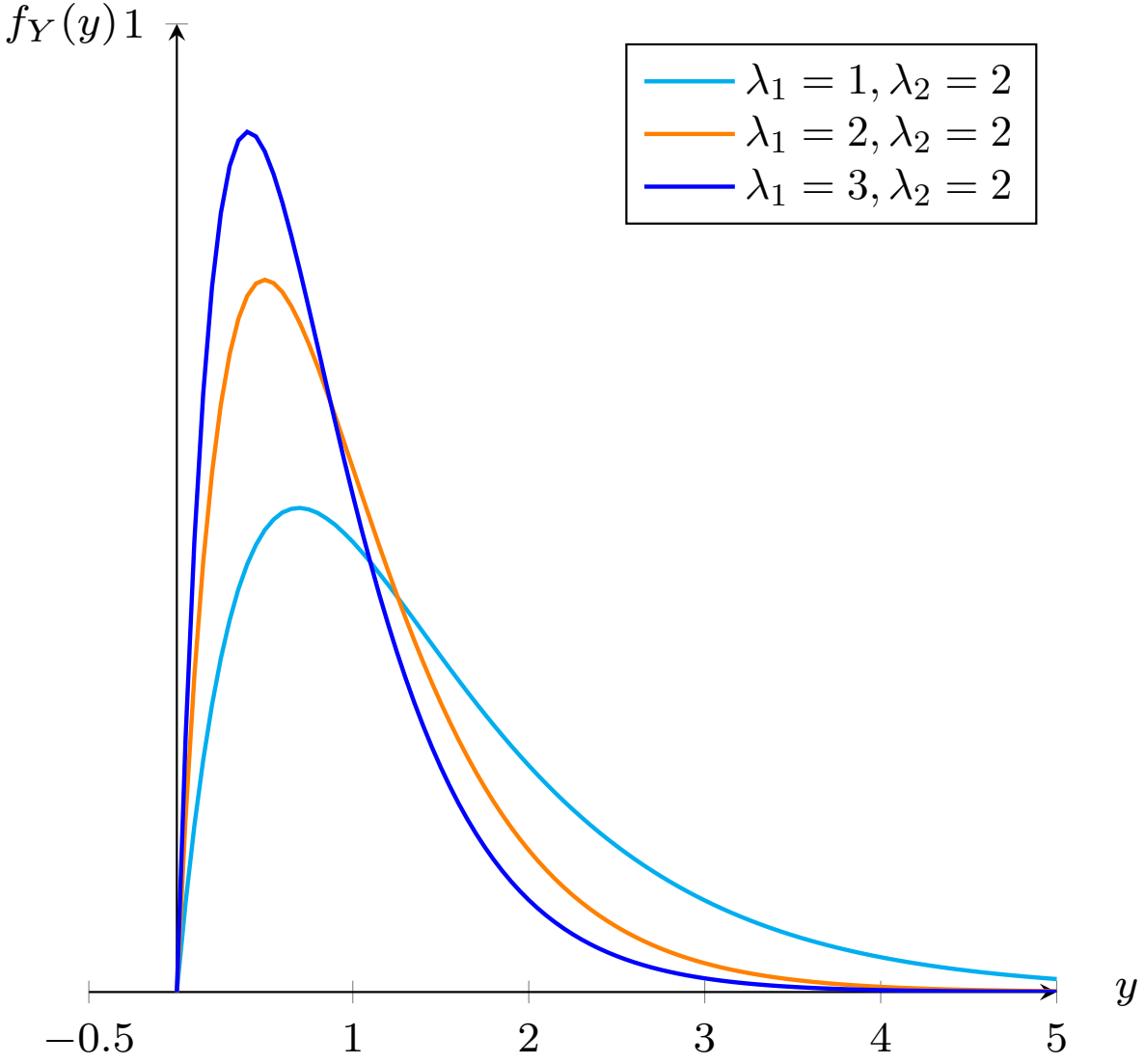
Example (Exponential random variables)

We then have two cases to consider:

$$f_Y(y) = \begin{cases} \dfrac{\lambda_1\lambda_2}{\lambda_1 - \lambda_2} e^{-\lambda_2 y}(1 - e^{-(\lambda_1-\lambda_2)y}) & \lambda_1 \neq \lambda_2 \\ \\ \lambda^2 y e^{-\lambda y} & \lambda_1 = \lambda_2 = \lambda, \text{ say.} \end{cases}$$

In the second case, we see that in fact Y has a gamma distribution with parameters $\alpha = 2$ and $\beta = 1/\lambda$.

Sums of random variables (cont.)



Sums of random variables (cont.)

Note

- We can extend the result to three or more independent random variables
$$Y = Y_1 + Y_2 + Y_3$$
by using the convolution formula recursively; we have
$$Y = (Y_1 + Y_2) + Y_3 = Y_{12} + Y_3$$
say, and so we would compute the pmf or pdf of $Y_{12} = Y_1 + Y_2$ first, and then convolve it with the pmf or pdf for Y_3.

- We can generalize to collections of random variables that are not independent.