Recap: Joint Distribution Function

- Given X, Y rv's on same probability space, joint distribution function: $F_{XY} : \Re^2 \to \Re$

$$F_{XY}(x, y) = P[X \leq x, Y \leq y]$$

- It satisfies
  1. $F_{XY}(-\infty, y) = F_{XY}(x, -\infty) = 0, \forall x, y;$
     $F_{XY}(\infty, \infty) = 1$
  2. F_{XY} is non-decreasing in each of its arguments
  3. F_{XY} is right continuous and has left-hand limits in each of its arguments
  4. For all $x_1 < x_2$ and $y_1 < y_2$

     $$F_{XY}(x_2, y_2) - F_{XY}(x_2, y_1) - F_{XY}(x_1, y_2) + F_{XY}(x_1, y_1) \geq 0$$

- Any $F : \Re^2 \to \Re$ satisfying the above would be a joint distribution function.

Recap: Joint Probability mass function

- $X \in \{x_1, x_2, \cdots\}$, $Y \in \{y_1, y_2, \cdots\}$
- The joint pmf: $f_{XY}(x, y) = P[X = x, Y = y]$.
- The joint pmf satisfies:
  - A1 $f_{XY}(x, y) \geq 0, \forall x, y$ and non-zero only for x_i, y_j pairs
  - A2 $\sum_i \sum_j f_{XY}(x_i, y_j) = 1$
- Given the joint pmf, we can get the joint df as

$$F_{XY}(x, y) = \sum_{\substack{i: \\ x_i \leq x}} \sum_{\substack{j: \\ y_j \leq y}} f_{XY}(x_i, y_j)$$

- Any $f_{XY} : \Re^2 \to [0,\ 1]$ satisfying $A1$ and $A2$ above is a joint pmf. (The F_{XY} satisfies all properties of df).
- Given the joint pmf, we can (in principle) compute the probability of any event involving the two discrete random variables.

$$P[(X, Y) \in B] = \sum_{\substack{i,j: \\ (x_i, y_j) \in B}} f_{XY}(x_i, y_j)$$

Recap: joint density

- Two cont rv X, Y have a joint density f_{XY} if

$$F_{XY}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(x', y')\, dy'\, dx', \;\; \forall x, y$$

  This implies $f_{XY}(x,y) = \frac{\partial^2}{\partial x\, \partial y} F_{XY}(x,y)$

- The joint density f_{XY} satisfies the following
  1. $f_{XY}(x,y) \geq 0, \;\; \forall x, y$
  2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x', y')\, dy'\, dx' = 1$
- Any function $f_{XY} : \Re^2 \rightarrow \Re$ satisfying the above two is a joint density function. (Then the above F_{XY} can be shown to be a joint df).
- We also have

$$P[x_1 \leq X \leq x_2, y_1 \leq Y \leq y_2] = \int_{x_1}^{x_2} \int_{y_1}^{y_2} f_{XY}\, dy\, dx$$

  and, $P[(X,Y) \in B] = \int_B f_{XY}(x,y)\, dx\, dy, \;\; \forall B \in \mathcal{B}^2$

Recap: Marginals

- Marginal distribution functions of X, Y are

$$F_X(x) = F_{XY}(x, \infty); \quad F_Y(y) = F_{XY}(\infty, y)$$

- X, Y discrete with joint pmf f_{XY}. The marginal pmfs are

$$f_X(x) = \sum_y f_{XY}(x, y); \quad f_Y(y) = \sum_x f_{XY}(x, y)$$

- If X, Y have joint pdf f_{XY} then the marginal pdf are

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

Recap: Conditional distributions

- Conditional distribution function of X conditioned on Y is

$$F_{X|Y}(x|y) = P[X \leq x | Y = y] \ \text{ when } Y \text{ is discrete}$$

$$F_{X|Y}(x|y) = \lim_{\delta \to 0} P[X \leq x | Y \in [y, y+\delta]] \ \text{ when } Y \text{ is continuous rv}$$

- This is well defined for all values that Y can assume.
- For each y, $F_{X|Y}(x|y)$ is a df in x.
- If X, Y have a joint density or if X is continuous and Y is discrete, $F_{X|Y}$ would have a density.

Recap: Contional density (or mass) fn

- Let X be a discrete random variable. Conditional mass fn of X conditioned on Y is

$$f_{X|Y}(x|y) = P[X = x|Y = y] \;\text{ if } Y \text{ is discrete}$$

$$f_{X|Y}(x|y) = \lim_{\delta \to 0} P[X = x|Y \in [y,\ y + \delta]\,] \;\text{ if } Y \text{ is continuous rv}$$

- This will be the mass function corresponding to the df $F_{X|Y}$.
- Let X be a continuous rv. Then we define conditional density $f_{X|Y}$ by

$$F_{X|Y}(x|y) = \int_{-\infty}^{x} f_{X|Y}(x'|y)\, dx'$$

This exists if X, Y have a joint density or when Y is discrete.

Recap

- When X, Y are both discrete or they have a joint density

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

- When X, Y are discrete or continuous (all four possibilities)

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

  Here $f_{X|Y}, f_X$ are densities when X is continuous and mass functions when X is discrete. Similarly for $f_{Y|X}, f_Y$
- The above relation gives rise to the total probability rules and Bayes rule for rv's

Recap

- If Y is discrete

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

- If X is continuous, the $f_X, f_{X|Y}$ are densities; If X is also discrete, they are mass functions
- If Y is continuous

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

- If X is also continuous, the $f_X, f_{X|Y}$ are densities; If X is discrete, they are mass functions (Where needed we assume the conditional density exists)

Recap: Bayes rule

- When X, Y are continuous or discrete (all four possibilities)

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

- This gives rise to Bayes rule:

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

- We need to interpret f_X or $f_{X|Y}$ as mass functions when X is discrete and as densities when X is a continuous and so on

Recap: Independent Random variables

- X and Y are said to be independent if events $[X \in B_1]$, $[Y \in B_2]$ are independent for all $B_1, B_2 \in \mathcal{B}$.
- X and Y are independent if and only if
  1. $F_{XY}(x, y) = F_X(x)\ F_Y(y)$
  2. $f_{XY}(x, y) = f_X(x)\ f_Y(y)$
- This also implies $F_{X|Y}(x|y) = F_X(x)$ and $f_{X|Y}(x|y) = f_X(x)$

Recap: More than two rv

- Everything is easily extended to multiple random variables.
- The joint distribution function of three rv's is

$$F_{XYZ}(x, y, z) = P[X \leq x, Y \leq y, Z \leq z]$$

- If all three are discrete then the joint mass function is

$$f_{XYZ}(x, y, z) = P[X = x, Y = y, Z = z]$$

- If they are continuous, they have a joint density if

$$F_{XYZ}(x, y, z) = \int_{-\infty}^{z} \int_{-\infty}^{y} \int_{-\infty}^{x} f_{XYZ}(x', y', z') \, dx' \, dy' \, dz'$$

Recap

- Joint mass function satisfies
  1. $f_{XYZ}(x, y, z) \geq 0$ and is non-zero only for countably many tuples.
  2. $\sum_{x,y,z} \; f_{XYZ}(x, y, z) = 1$
- Joint density satisfies
  1. $f_{XYZ}(x, y, z) \geq 0$
  2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \; f_{XYZ}(x, y, z) \; dx \; dy \; dz = 1$

Recap

- Now we get many different marginals:

$$F_{XY}(x, y) = F_{XYZ}(x, y, \infty); \;\; F_Z(z) = F_{XYZ}(\infty, \infty, z) \quad \text{and so on}$$

- Any marginal density is a joint density of a subset of these rv's and we obtain it by integrating the (full) joint density with respect to the remaining variables.

$$\begin{aligned} f_{YZ}(y, z) &= \int_{-\infty}^{\infty} f_{XYZ}(x, y, z)\, dx; \\ f_X(x) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XYZ}(x, y, z)\, dy\, dz \end{aligned}$$

- We obtain the marginal mass functions for a subset of the rv's also similarly where we sum over the remaining variables.

Recap

- Like in case of marginals, there are different types of conditional distributions now.
- We can always define conditional distribution functions like

$$\begin{aligned} F_{XY|Z}(x, y|z) &= P[X \leq x, Y \leq y | Z = z] \\ F_{X|YZ}(x|y, z) &= P[X \leq x | Y = y, Z = z] \\ F_{XY|ZW}(x, y|z, w) &= P[X \leq x, Y \leq y | Z = z, W = w] \end{aligned}$$

- In all such cases, if the conditioning random variables are continuous, we define the above as a limit.
- For example when Z is continuous

$$F_{XY|Z}(x, y|z) = \lim_{\delta \to 0} P[X \leq x, Y \leq y | Z \in [z, z + \delta]]$$

Recap

- If X, Y, Z are all discrete then, all conditional mass functions are defined by appropriate conditional probabilities.
- Conditional densities are the densities corresponding to conditional distributions.
- In all cases, the following type relations hold

$$\begin{aligned}
f_{XY|Z}(x, y|z) &= \frac{f_{XYZ}(x, y, z)}{f_Z(z)} \\
f_{X|YZ}(x|y, z) &= \frac{f_{XYZ}(x, y, z)}{f_{YZ}(y, z)} \\
f_{XYZ}(x, y, z) &= f_{Z|YX}(z|y, x) f_{Y|X}(y|x) f_X(x)
\end{aligned}$$

Recap

- The notation when we consider n random variables is the following
- Let $\mathbf{X} = (X_1, X_2, \cdots, X_n)$. We can think of it as a mapping, $\quad \mathbf{X} : \Omega \to \Re^n$.
- We can write the joint distribution function as

$$F_{\mathbf{X}}(\mathbf{x}) = P[\mathbf{X} \leq \mathbf{x}] = P[X_i \leq x_i,\ i = 1, \cdots, n]$$

- We represent by $f_{\mathbf{X}}(\mathbf{x})$ the joint density or mass function.
- We use similar notation for marginal and conditional distributions

Recap: Independence of multiple random variables

- Random variables $X_1, X_2, \cdots, X_n$ are said to be independent if the the events $[X_i \in B_i],\ i = 1, \cdots, n$ are independent.
- Independence implies that the marginals would determine the joint distribution.
- If $X_1, X_2, \cdots, X_n$ are independent and if each X_i has the same distribution, they are said to be **independent and identically distributed** or **iid** random variables.

Recap: Functions of two random variables

- Let X, Y be random variables on the same probability space.
- Let $g : \Re^2 \to \Re$.
- Let $Z = g(X, Y)$. Then Z is a rv (assuming g is 'nice')
- We can determine distribution of Z from the joint distribution of X, Y

$$F_Z(z) = P[Z \leq z] = P[g(X, Y) \leq z]$$

- For example, if X, Y are discrete, then

$$f_Z(z) = P[Z = z] = P[g(X, Y) = z] = \sum_{\substack{x_i, y_j : \\ g(x_i, y_j) = z}} f_{XY}(x_i, y_j)$$

Recap: Functions of multiple rv's

- Given $X_1, \cdots, X_n$, random variables on the same probability space, $Z = g(X_1, \cdots, X_n)$ is a rv (if $g : \Re^n \to \Re$ is borel measurable).
- We can determine distribution of Z from the joint distribution of all X_i

$$F_Z(z) = P[Z \leq z] = P[g(X_1, \cdots, X_n) \leq z]$$

Recap: max of n rv's

- Let $X_1, \cdots, X_n$ be independent and $Z = \max(X_1, \cdots, X_n)$

$$\begin{aligned} F_Z(z) &= P[\max(X_1, \cdots, X_n) \leq z] \\ &= P[X_i \leq z, \ i = 1, \cdots, n] \\ &= \prod_{i=1}^{n} F_{X_i}(z) \\ &= (F(z))^n, \quad \text{if they are iid} \end{aligned}$$

Recap: min of n rv's

- Let $X_1, \cdots, X_n$ be independent and $Z = \min(X_1, \cdots, X_n)$
- We have $P[Z > z] = P[X_i > z, \; i = 1, \cdots n]$.
- This gives

$$
\begin{aligned}
F_Z(z) &= 1 - \prod_{i=1}^{n} (1 - F_{X_i}(z)) \\
&= 1 - (1 - F(z))^n, \quad \text{if they are iid}
\end{aligned}
$$

Joint distribution of max and min

- X, Y iid with df F and density f
  $Z = \max(X, Y)$ and $W = \min(X, Y)$.
- We want joint distribution function of Z and W.
- We can use the following

$$P[Z \leq z] = P[Z \leq z, W \leq w] + P[Z \leq z, W > w]$$

$$\begin{aligned} P[Z \leq z, W > w] &= P[w < X, Y \leq z] = (F(z) - F(w))^2 \\ P[Z \leq z] &= P[X \leq z, Y \leq z] = (F(z))^2 \end{aligned}$$

- So, we get F_{ZW} as

$$\begin{aligned} F_{ZW}(z, w) &= P[Z \leq z, W \leq w] \\ &= P[Z \leq z] - P[Z \leq z, W > w] \\ &= (F(z))^2 - (F(z) - F(w))^2 \end{aligned}$$

- Is this correct for all values of z, w?

- We have $P[w < X, Y \leq z] = (F(z) - F(w))^2$ only when $w \leq z$.
- Otherwise it is zero.
- Hence we get F_{ZW} as

$$F_{ZW}(z,w) = \begin{cases} (F(z))^2 & \text{if} \quad w > z \\ (F(z))^2 - (F(z) - F(w))^2 & \text{if} \quad w \leq z \end{cases}$$

- We can get joint density of Z, W as

$$\begin{aligned} f_{ZW}(z,w) &= \frac{\partial^2}{\partial z \, \partial w} F_{ZW}(z,w) \\ &= 2f(z)f(w), \quad w \leq z \end{aligned}$$

- Let X, Y be iid uniform over $(0,\ 1)$.
- Define $Z = \max(X, Y)$ and $W = \min(X, Y)$.
- Then the joint density of Z, W is

$$\begin{aligned} f_{ZW}(z,w) &= 2f(z)f(w), \quad w \leq z \\ &= 2, \quad 0 < w \leq z < 1 \end{aligned}$$

Order Statistics

- Let $X_1, \cdots, X_n$ be iid with density f.
- Let $X_{(k)}$ denote the k^{th} smallest of these.
- That is, $X_{(k)} = g_k(X_1, \cdots, X_n)$ where $g_k : \Re^n \rightarrow \Re$ and the value of $g_k(x_1, \cdots, x_n)$ is the k^{th} smallest of the numbers $x_1, \cdots, x_n$.
- $X_{(1)} = \min(X_1, \cdots, X_n), \quad X_{(n)} = \max(X_1, \cdots, X_n)$
- The joint distribution of $X_{(1)}, \cdots X_{(n)}$ is called the order statistics.
- Earlier, we calculated the order statistics for the case $n = 2$.
- It can be shown that

$$f_{X_{(1)} \cdots X_{(n)}}(x_1, \cdots x_n) = n! \prod_{i=1}^{n} f(x_i), \quad x_1 < x_2 < \cdots < x_n$$

Marginal distributions of $X_{(k)}$

- Let $X_1, \cdots, X_n$ be iid with df F and density f.
- Let $X_{(k)}$ denote the k^{th} smallest of these.
- We want the distribution of $X_{(k)}$.
- The event $[X_{(k)} \leq y]$ is:
  "at least k of these are less than or equal to y"
- We want probability of this event.

Marginal distributions of $X_{(k)}$

- $X_1, \cdots, X_n$ iid with df F and density f.
- $P[X_i \leq y] = F(y)$ for any i and y.
- Since they are independent, we have, e.g.,

$$P[X_1 \leq y, X_2 > y, X_3 \leq y] = (F(y))^2(1 - F(y))$$

- Hence, probability that exactly k of these n random variables are less than or equal to y is ${}^nC_k(F(y))^k(1 - F(y))^{n-k}$
- Hence we get

$$F_{X_{(k)}}(y) = \sum_{j=k}^{n} {}^nC_j(F(y))^j(1 - F(y))^{n-j}$$

We can get the density by differentiating this.

Sum of two discrete rv's

- Let $X, Y \in \{0, 1, \cdots\}$
- Let $Z = X + Y$. Then we have

$$\begin{aligned} f_Z(z) &= P[X+Y=z] = \sum_{\substack{x,y: \\ x+y=z}} P[X=x, Y=y] \\ &= \sum_{k=0}^{z} P[X=k, Y=z-k] \\ &= \sum_{k=0}^{z} f_{XY}(k, z-k) \end{aligned}$$

- Now suppose X, Y are independent. Then

$$f_Z(z) = \sum_{k=0}^{z} f_X(k) f_Y(z-k)$$

- Now suppose X, Y are independent Poisson with parameters λ_1, λ_2. And, $Z = X + Y$.

$$
\begin{aligned}
f_Z(z) &= \sum_{k=0}^{z} f_X(k) f_Y(z-k) \\
&= \sum_{k=0}^{z} \frac{\lambda_1^k}{k!} e^{-\lambda_1} \; \frac{\lambda_2^{z-k}}{(z-k)!} e^{-\lambda_2} \\
&= e^{-(\lambda_1+\lambda_2)} \; \frac{1}{z!} \sum_{k=0}^{z} \frac{z!}{k!(z-k)!} \; \lambda_1^k \; \lambda_2^{z-k} \\
&= e^{-(\lambda_1+\lambda_2)} \; \frac{1}{z!} (\lambda_1 + \lambda_2)^z
\end{aligned}
$$

- Z is Poisson with parameter $\lambda_1 + \lambda_2$

Sum of two continuous rv

- Let X, Y have a joint density f_{XY}. Let $Z = X + Y$

$$
\begin{aligned}
F_Z(z) &= P[Z \leq z] = P[X + Y \leq z] \\
&= \int\int_{\{(x,y): x+y \leq z\}} f_{XY}(x, y)\, dy\, dx \\
&= \int_{x=-\infty}^{\infty} \int_{y=-\infty}^{z-x} f_{XY}(x, y)\, dy\, dx
\end{aligned}
$$

change variable y to t: $t = x + y$

$$dt = dy; \quad y = z - x \;\Rightarrow\; t = z$$

$$
\begin{aligned}
&= \int_{x=-\infty}^{\infty} \int_{t=-\infty}^{z} f_{XY}(x, t - x)\, dt\, dx \\
&= \int_{-\infty}^{z} \left(\int_{-\infty}^{\infty} f_{XY}(x, t - x)\, dx \right) dt
\end{aligned}
$$

- This gives us the density of Z

- X, Y have joint density f_{XY}. $Z = X + Y$. Then

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

- Now suppose X and Y are independent. Then

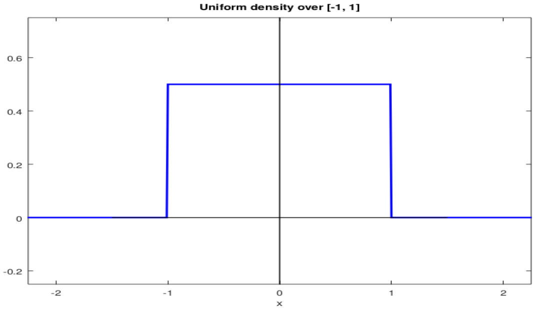
$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(z - x)\, dx$$

Density of sum of independent random variables is the convolution of their densities.

$$f_{X+Y} = f_X * f_Y \quad \text{(Convolution)}$$

Distribution of sum of iid uniform rv's

- Suppose X, Y are iid uniform over $(-1,\ 1)$.
- let $Z = X + Y$. We want f_Z.
- The density of X, Y is



- f_Z is convolution of this density with itself.

- $f_X(x) = 0.5, \quad -1 < x < 1.$ f_Y is also same
- Note that Z takes values in $[-2,\ 2]$

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(t)\ f_Y(z-t)\ dt$$

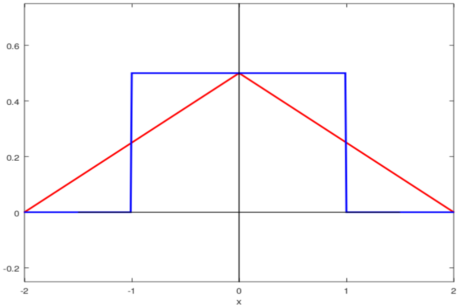
- For the integrand to be non-zero we need
  - $-1 < t < 1 \quad \Rightarrow \quad t < 1, \quad t > -1$
  - $-1 < z - t < 1 \ \Rightarrow \quad t < z + 1, \quad t > z - 1$
  - Hence we need:
    $t < \min(1, z+1), \quad t > \max(-1, z-1)$
  - Hence, for $z < 0$, we need $-1 < t < z + 1$
    and, for $z \geq 0$ we need $z - 1 < t < 1$
- Thus we get

$$f_Z(z) = \begin{cases} \int_{-1}^{z+1} \frac{1}{4} dt = \frac{z+2}{4} & \text{if } \ -2 \leq z < 0 \\ \int_{z-1}^{1} \frac{1}{4} dt = \frac{2-z}{4} & \text{if } \ 2 \geq z \geq 0 \end{cases}$$

- Thus, the density of sum of two ind rv's that are uniform over $(-1,\ 1)$ is

$$f_Z(z) = \begin{cases} \frac{z+2}{4} & \text{if } \ -2 < z < 0 \\ \frac{2-z}{4} & \text{if } \ 0 < z < 2 \end{cases}$$

- This is a triangle with vertices $(-2,0), (0,0.5), (2,0)$



Independence of functions of random variable

- Suppose X and Y are independent.
- Then $g(X)$ and $h(Y)$ are independent
- This is because $[g(X) \in B_1] = [X \in \tilde{B}_1]$ for some Borel set, $\tilde{B}_1$ and similarly $[h(Y) \in B_2] = [Y \in \tilde{B}_2]$
- Hence, $[g(X) \in B_1]$ and $[h(Y) \in B_2]$ are independent.

Independence of functions of random variable

- This is easily generalized to functions of multiple random variables.
- If $\mathbf{X}, \mathbf{Y}$ are vector random variables (or random vectors), independence implies $[\mathbf{X} \in B_1]$ is independent of $[\mathbf{Y} \in B_2]$ for all borel sets B_1, B_2 (in appropriate spaces).
- Then $g(\mathbf{X})$ would be independent of $h(\mathbf{Y})$.
- That is, suppose $X_1, \cdots, X_m, Y_1, \cdots, Y_n$ are independent.
- Then, $g(X_1, \cdots, X_m)$ is independent of $h(Y_1, \cdots, Y_n)$.

- Let X_1, X_2, X_3 be independent continuous rv
- $Z = X_1 + X_2 + X_3$.
- Can we find density of Z?
- Let $W = X_1 + X_2$.
- Then $Z = W + X_3$ and W and X_3 are independent.
- Exercise for you: Find density of $X_1 + X_2 + X_3$ where X_1, X_2, X_3 are iid uniform over $(0,\ 1)$.

- Suppose X, Y are iid exponential rv's.

$$f_X(x) = \lambda\, e^{-\lambda x}, \;\; x > 0$$

- Let $Z = X + Y$. Then, density of Z is

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_X(x)\, f_Y(z-x)\, dx \\ &= \int_0^z \lambda\, e^{-\lambda x}\, \lambda\, e^{-\lambda(z-x)}\, dx \\ &= \lambda^2\, e^{-\lambda z} \int_0^z dx = \lambda^2\, z\, e^{-\lambda z} \end{aligned}$$

- Thus, sum of independent exponential random variables has gamma distribution:

$$f_Z(z) = \lambda z\, \lambda e^{-\lambda z}, \;\; z > 0$$

Sum of independent gamma rv

- Gamma density with parameters $\alpha > 0$ and $\lambda > 0$ is given by
$$f(x) = \frac{1}{\Gamma(\alpha)} \, \lambda^{\alpha} \, x^{\alpha-1} \, e^{-\lambda x}, \quad x > 0$$
We will call this $Gamma(\alpha, \lambda)$.
- The α is called the shape parameter and λ is called the rate parameter.
- For $\alpha = 1$ this is the exponential density.
- Let $X \sim Gamma(\alpha_1, \lambda)$, $Y \sim Gamma(\alpha_2, \lambda)$. Suppose X, Y are independent.
- Let $Z = X + Y$. Then $Z \sim Gamma(\alpha_1 + \alpha_2, \lambda)$.

$$f_Z(z) \;=\; \int_{-\infty}^{\infty} f_X(x)\, f_Y(z-x)\, dx$$

$$= \int_0^z \frac{1}{\Gamma(\alpha_1)}\, \lambda^{\alpha_1}\, x^{\alpha_1 - 1}\, e^{-\lambda x} \frac{1}{\Gamma(\alpha_2)}\, \lambda^{\alpha_2}\, (z-x)^{\alpha_2 - 1}\, e^{-\lambda(z-x)}\, dx$$

$$= \frac{\lambda^{\alpha_1+\alpha_2}\, e^{-\lambda z}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \int_0^z z^{\alpha_1-1} \left(\frac{x}{z}\right)^{\alpha_1 - 1} z^{\alpha_2 - 1} \left(1 - \frac{x}{z}\right)^{\alpha_2 - 1}\, dx$$

change the variable: $t = \frac{x}{z}$ $(\Rightarrow \;\; z^{-1} dx = dt)$

$$= \frac{\lambda^{\alpha_1+\alpha_2}\, e^{-\lambda z}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\, z^{\alpha_+\alpha_2-1} \int_0^1 t^{\alpha_1-1}\, (1-t)^{\alpha_2-1}\, dt$$

$$= \frac{1}{\Gamma(\alpha_1+\alpha_2)}\, \lambda^{\alpha_1+\alpha_2}\, z^{\alpha_1+\alpha_2-1}\, e^{-\lambda z}$$

Because

$$\int_0^1 t^{\alpha_1-1}\, (1-t)^{\alpha_2-1}\, dt = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)}$$

- If X, Y are independent gamma random variables then $X + Y$ also has gamma distribution.
- If $X \sim Gamma(\alpha_1, \lambda)$, and $Y \sim Gamma(\alpha_2, \lambda)$, then $X + Y \sim Gamma(\alpha_1 + \alpha_2, \lambda)$.

Sum of independent Gaussians

- Sum of independent Gaussians random variables is a Gaussian rv
- If $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$ and X, Y are independent, then $X + Y \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$
- Exercise for you: Show that sum of independent Gaussian random variables has gaussian density.
- The algebra is a little involved.
- First take the two gaussians to be zero-mean.
- There is a calculation trick that is often useful with Gaussian density

A Calculation Trick

$$
\begin{aligned}
I &= \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2K}\left[x^2 - 2bx + c\right]\right)\, dx \\
&= \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2K}\left[(x-b)^2 + c - b^2\right]\right)\, dx \\
&= \int_{-\infty}^{\infty} \exp\left(-\frac{(x-b)^2}{2K}\right)\, \exp\left(-\frac{(c-b^2)}{2K}\right)\, dx \\
&= \exp\left(-\frac{(c-b^2)}{2K}\right)\sqrt{2\pi K}
\end{aligned}
$$

because

$$
\frac{1}{\sqrt{2\pi K}} \int_{-\infty}^{\infty} \exp\left(-\frac{(x-b)^2}{2K}\right)\, dx = 1
$$

- We next look at a general theorem that is quite useful in dealing with functions of multiple random variables.
- This result is only for continuous random variables.

- Let $X_1, \cdots, X_n$ be continuous random variables with joint density $f_{X_1 \cdots X_n}$. We define $Y_1, \cdots Y_n$ by

$$Y_1 = g_1(X_1, \cdots, X_n) \quad \cdots \quad Y_n = g_n(X_1, \cdots, X_n)$$

  We think of g_i as components of $g : \Re^n \to \Re^n$.
- We assume g is continuous with continuous first partials and is invertible.
- Let h be the inverse of g. That is

$$X_1 = h_1(Y_1, \cdots, Y_n) \quad \cdots \quad X_n = h_n(Y_1, \cdots, Y_n)$$

- Each of g_i, h_i are $\Re^n \to \Re$ functions and we can write them as

$$y_i = g_i(x_1, \cdots, x_n); \quad \cdots \quad x_i = h_i(y_1, \cdots, y_n)$$

  We denote the partial derivatives of these functions by $\frac{\partial x_i}{\partial y_j}$ etc.

- The jacobian of the inverse transformation is

$$J = \frac{\partial(x_1, \cdots, x_n)}{\partial(y_1, \cdots, y_n)} = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \cdots & \frac{\partial x_2}{\partial y_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \frac{\partial x_n}{\partial y_2} & \cdots & \frac{\partial x_n}{\partial y_n} \end{vmatrix}$$

- We assume that J is non-zero in the range of the transformation
- **Theorem**: Under the above conditions, we have

$$f_{Y_1 \cdots Y_n}(y_1, \cdots, y_n) = |J| f_{X_1 \cdots X_n}\left(h_1(y_1, \cdots, y_n), \cdots, h_n(y_1, \cdots, y_n)\right)$$

Or, more compactly, $\quad f_{\mathbf{Y}}(\mathbf{y}) = |J| f_{\mathbf{X}}(h(\mathbf{y}))$

- Let X_1, X_2 have a joint density, $f_{\mathbf{X}}$. Consider

$$\begin{aligned} Y_1 &= g_1(X_1, X_2) = X_1 + X_2 \quad (g_1(a,b) = a+b) \\ Y_2 &= g_2(X_1, X_2) = X_1 - X_2 \quad (g_2(a,b) = a-b) \end{aligned}$$

  This transformation is invertible

$$\begin{aligned} X_1 &= h_1(Y_1, Y_2) = \frac{Y_1 + Y_2}{2} \quad (h_1(a,b) = (a+b)/2) \\ X_2 &= h_2(Y_1, Y_2) = \frac{Y_1 - Y_2}{2} \quad (h_2(a,b) = (a-b)/2) \end{aligned}$$

  The jacobian is: $\begin{vmatrix} 0.5 & 0.5 \\ 0.5 & -0.5 \end{vmatrix} = -0.5.$

- This gives: $f_{Y_1 Y_2}(y_1, y_2) = 0.5 \; f_{X_1 X_2}\left(\frac{y_1+y_2}{2}, \; \frac{y_1-y_2}{2}\right)$

Proof of Theorem

- Let $B = (-\infty,\ y_1] \times \cdots \times (-\infty,\ y_n] \subset \Re^n$. Then

$$\begin{aligned} F_{\mathbf{Y}}(\mathbf{y}) &= F_{Y_1 \cdots Y_n}(y_1, \cdots y_n) = P[Y_i \leq y_i,\ i = 1, \cdots, n] \\ &= \int_B f_{Y_1 \cdots Y_n}(y_1', \cdots, y_n')\, dy_1' \ \cdots \ dy_n' \end{aligned}$$

- Define

$$\begin{aligned} g^{-1}(B) &= \{(x_1, \cdots, x_n) \in \Re^n \ : \ g(x_1, \cdots, x_n) \in B\} \\ &= \{(x_1, \cdots, x_n) \in \Re^n \ : \ g_i(x_1, \cdots, x_n) \leq y_i,\ i = 1 \cdots n\} \end{aligned}$$

- Then we have

$$\begin{aligned} F_{Y_1 \cdots Y_n}(y_1, \cdots y_n) &= P[g_i(X_1, \cdots, X_n) \leq y_i,\ i = 1, \cdots n] \\ &= \int_{g^{-1}(B)} f_{X_1 \cdots X_n}(x_1', \cdots, x_n')\, dx_1' \ \cdots \ dx_n' \end{aligned}$$

Proof of Theorem

- $B = (-\infty,\ y_1] \times \cdots \times (-\infty,\ y_n]$.
- $g^{-1}(B) = \{(x_1, \cdots, x_n) \in \Re^n \ :\ g(x_1, \cdots, x_n) \in B\}$

$$
\begin{aligned}
F_{\mathbf{Y}}(y_1, \cdots, y_n) &= P[g_i(X_1, \cdots, X_n) \le y_i,\ i = 1, \cdots, n] \\
&= \int_{g^{-1}(B)} f_{X_1 \cdots X_n}(x_1', \cdots, x_n')\, dx_1' \ \cdots\ dx_n'
\end{aligned}
$$

change variables: $y_i' = g_i(x_1', \cdots, x_n'), i = 1, \cdots n$

$$(x_1', \cdots x_n') \in g^{-1}(B) \ \Rightarrow\ (y_1', \cdots, y_n') \in B$$

$$x_i' = h_i(y_1', \cdots, y_n'), \quad dx_1' \cdots dx_n' = |J| dy_1' \cdots dy_n'$$

$$F_{\mathbf{Y}}(y_1, \cdots, y_n) = \int_B f_{X_1 \cdots X_n}(h_1(\mathbf{y}'), \cdots, h_n(\mathbf{y}'))\ |J| dy_1' \cdots dy_n'$$

$$\Rightarrow\ f_{Y_1 \cdots Y_n}(y_1, \cdots, y_n) = f_{X_1 \cdots X_n}(h_1(\mathbf{y}), \cdots, h_n(\mathbf{y}))\ |J|$$

- $X_1, \cdots X_n$ are continuous rv with joint density

$$Y_1 = g_1(X_1, \cdots, X_n) \quad \cdots \quad Y_n = g_n(X_1, \cdots, X_n)$$

- The transformation is continuous with continuous first partials and is invertible and

$$X_1 = h_1(Y_1, \cdots, Y_n) \quad \cdots \quad X_n = h_n(Y_1, \cdots, Y_n)$$

- We assume the Jacobian of the inverse transform, J, is non-zero
- Then the density of $\mathbf{Y}$ is

$$f_{Y_1 \cdots Y_n}(y_1, \cdots, y_n) = |J| f_{X_1 \cdots X_n}\left(h_1(y_1, \cdots, y_n), \cdots, h_n(y_1, \cdots, y_n)\right)$$

- Called multidimensional change of variable formula

- Let X, Y have joint density f_{XY}. Let $Z = X + Y$.
- We want to find f_Z using the theorem.
- To use the theorem, we need an invertible transformation of $\Re^2$ onto $\Re^2$ of which one component is $x + y$.
- Take $Z = X + Y$ and $W = X - Y$. This is invertible.
- $X = (Z + W)/2$ and $Y = (Z - W)/2$. The Jacobian is

$$J = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}$$

- Hence we get

$$f_{ZW}(z, w) = \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right)$$

- Now we get density of Z as

$$f_Z(z) = \int_{-\infty}^{\infty} \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right)\ dw$$