Marginal Probability Distribution

To compute $p_X(k)$, we sum $p_{XY}(i,j)$ over pairs of i and j where $i = k$:

$$
\begin{aligned}
p_X(k) &= \sum_{\{i,j \mid i=k\}} p_{XY}(k,j) \\
&= \sum_{j=-\infty}^{\infty} p_{XY}(k,j).
\end{aligned}
$$

Sum of Discrete r.v.'s

Let Z be a discrete random variable equal to the sum of the discrete random variables X and Y. To compute $p_Z(k)$, we sum $p_{XY}(i,j)$ over pairs of i and j where $i+j=k$:

$$p_Z(k) = \sum_{\{i,j \mid i+j=k\}} p_{XY}(i,j).$$

Observe that $i+j=k$ iff $j=k-i$. It follows that

$$p_Z(k) = \sum_{i=-\infty}^{\infty} p_{XY}(i,k-i)$$

If X and Y are statistically independent, then $p_{XY}(i,j) = p_X(i)p_Y(j)$. It follows that

$$p_Z(k) = \sum_{i=-\infty}^{\infty} p_X(i)p_Y(k-i).$$

Sum of Discrete r.v.'s (contd.)

The *discrete convolution* of f and g, written $f * g$, is defined to be:

$$\{f * g\}(k) = \sum_{i=-\infty}^{\infty} f(i)g(k-i).$$

Accordingly, if $Z = X + Y$, then

$$p_Z = p_X * p_Y.$$

Sum of Discrete r.v.'s (contd.)

Since $i+j=k$ iff $i=k-j$, we could just as easily have written $p_Z(k)$ as follows:

$$
\begin{aligned}
p_Z(k) &= \sum_{j=-\infty}^{\infty} p_{XY}(k-j, j) \\
&= \sum_{j=-\infty}^{\infty} p_X(k-j) p_Y(j).
\end{aligned}
$$

It follows that

$$p_X * p_Y = p_Y * p_X$$

and that convolution (like addition) is *commutative*.

Difference of Discrete r.v.'s

Let Z be a discrete random variable equal to the difference of the discrete random variables X and Y. To compute $p_Z(k)$, we sum $p_{XY}(i,j)$ over pairs of i and j where $i-j=k$:

$$p_Z(k) = \sum_{\{i,j \mid i-j=k\}} p_{XY}(i,j).$$

Observe that $i-j=k$ iff $j=i-k$. It follows that

$$p_Z(k) = \sum_{i=-\infty}^{\infty} p_{XY}(i,i-k).$$

If X and Y are statistically independent, then $p_{XY}(i,j) = p_X(i)p_Y(j)$. It follows that

$$p_Z(k) = \sum_{i=-\infty}^{\infty} p_X(i)p_Y(i-k).$$

Difference of Discrete r.v.'s (contd).

The *discrete correlation* of f and g, is defined to be:

$$\{f * g(-(.))\}(k) = \sum_{i=-\infty}^{\infty} f(i)g(i-k).$$

Marginal Probability Density

To compute $f_X(z)$, we integrate $f_{XY}(x, y)$ along the line $x = z$:

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

Sum of Continuous r.v.'s

Let Z be a continuous random variable equal to the sum of the continuous random variables, X and Y. To compute $f_Z(z)$, we integrate $f_{XY}(x,y)$ along the line $x+y=z$ or $y=z-x$:

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

If X and Y are statistically independent, then $f_{XY}(x,y)=f_X(x)f_Y(y)$. It follows that

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

Sum of Continuous r.v.'s (contd.)

The *convolution* of f and g, written $f*g$, is defined to be

$$\{f*g\}(v)=\int_{-\infty}^{\infty}f(u)g(v-u)du.$$

Accordingly, if $Z=X+Y$, then

$$f_Z=f_X*f_Y=f_Y*f_X.$$

Difference of Continuous r.v.'s

Let Z be a continuous random variable equal to the difference of the continuous random variables, X and Y. To compute $f_Z(z)$, we integrate $f_{XY}(x,y)$ along the line where $x - y = z$ or $y = x - z$:

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

If X and Y are statistically independent, then $f_{XY}(x,y) = f_X(x)f_Y(y)$. It follows that

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

Difference of Continuous r.v.'s (contd.)

The *correlation* of f and g, is defined to be

$$\{f * g(-(.))\}(v) = \int_{-\infty}^{\infty} f(u)g(u-v)du$$

Accordingly, if $Z = X - Y$, then

$$f_Z(z) = \{f_X * f_Y(-(.))\}(z)$$

Expected Value of Sum and Product

The expected value of $X+Y$ is

$$\begin{aligned}\langle X+Y\rangle &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)\,f_{XY}(x,y)dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x\,f_{XY}(x,y)dxdy+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}y\,f_{XY}(x,y)dxdy \\ &= \int_{-\infty}^{\infty}x\int_{-\infty}^{\infty}f_{XY}(x,y)dydx+\int_{-\infty}^{\infty}y\int_{-\infty}^{\infty}f_{XY}(x,y)dxdy \\ &= \int_{-\infty}^{\infty}xf_X(x)dx+\int_{-\infty}^{\infty}y\,f_Y(y)dy \\ &= \langle X\rangle+\langle Y\rangle\,.\end{aligned}$$

When X and Y are independent, the expected value of XY is

$$\begin{aligned}\langle XY\rangle &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x\,y\,f_{XY}(x,y)dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x\,y\,f_X(x)\,f_Y(y)dxdy \\ &= \int_{-\infty}^{\infty}x\,f_X(x)dx\int_{-\infty}^{\infty}y\,f_Y(y)dy \\ &= \langle X\rangle\,\langle Y\rangle\,.\end{aligned}$$

Law of Large Numbers

Let $x_1 \ldots x_N$ be samples of a r.v., X. It follows that $x_1 \ldots x_N$ are independent, identically distributed (i.i.d.) random variables. The *Law of Large Numbers* states that

$$\lim_{N \to \infty} \frac{\sum_{i=1}^{N} x_i}{N} = \mu_X$$

i.e., the average of an infinite number of samples equals the expected value. Now define a new random variable Y:

$$Y = (X - \mu_X)^2$$

where $y_i = (x_i - \mu_X)^2$. Observe that N samples of $y_1 \ldots y_N$ are also i.i.d. random variables. Consequently,

$$\lim_{N \to \infty} \frac{\sum_{i=1}^{N} y_i}{N} = \mu_Y = \lim_{N \to \infty} \frac{\sum_{i=1}^{N} (x_i - \mu_X)^2}{N} = \sigma_X^2.$$

Variance of Sum of Continuous r.v.'s

Let X and Y be two independent zero mean r.v.'s and let $Z = X + Y$, then

$$
\begin{aligned}
\sigma_Z^2 &= \lim_{N\to\infty} \frac{\sum_{i=1}^N z_i^2}{N} \\
&= \lim_{N\to\infty} \frac{\sum_{i=1}^N (x_i + y_i)^2}{N} \\
&= \lim_{N\to\infty} \frac{\sum_{i=1}^N \left(x_i^2 + 2\,x_i\,y_i + y_i^2\right)}{N} \\
&= \lim_{N\to\infty} \left[\frac{\sum_{i=1}^N x_i^2}{N} + \underbrace{\frac{\sum_{i=1}^N 2\,x_i\,y_i}{N}}_{0} + \frac{\sum_{i=1}^N y_i^2}{N} \right] \\
&= \lim_{N\to\infty} \frac{\sum_{i=1}^N x_i^2}{N} + \lim_{N\to\infty} \frac{\sum_{i=1}^N y_i^2}{N} \\
&= \sigma_X^2 + \sigma_Y^2.
\end{aligned}
$$

Central Limit Theorem

If $x_1 \ldots x_N$ are independent samples of a random variable X and if

$$\lim_{N \to \infty} \frac{\sum_{i=1}^{N} x_i}{N} = \mu_X = 0$$

and

$$\lim_{N \to \infty} \frac{\sum_{i=1}^{N} x_i^2}{N} = \sigma_X^2 = 1$$

then

$$F_Z(a) = \lim_{N \to \infty} P\left(\frac{\sum_{i=1}^{N} x_i}{\sqrt{N}} \leq a\right) =$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-z^2/2} dz.$$

In other words, the distribution of the sum Z of an infinite number of samples of a random variable X with zero mean and unit variance is normal.

Central Limit Theorem (contd.)

Assume that there exists a p.d.f. f_Z characterizing the distribution of the sum Z. Now define

$$\begin{aligned} Z_{\text{even}} &= \lim_{N\to\infty} \frac{\sum_{i=1}^{N} x_{2i}}{\sqrt{N}} \\ Z_{\text{odd}} &= \lim_{N\to\infty} \frac{\sum_{i=1}^{N} x_{2i+1}}{\sqrt{N}} \\ Z_{\text{both}} &= \lim_{N\to\infty} \frac{\sum_{i=1}^{2N} x_{i}}{\sqrt{2N}}. \end{aligned}$$

Since Z_{even}, Z_{odd}, Z_{both} and Z are all sums of infinite numbers of samples of X, it follows that

$$P(Z_{\text{even}} \leq a) = P(Z_{\text{odd}} \leq a) = P(Z_{\text{both}} \leq a) = F_Z(a)$$

where

$$F_Z(a) = \int_0^a f_Z(z)dz.$$

Central Limit Theorem (contd.)

Now, observe that

$$\begin{aligned}
\lim_{N\to\infty} \frac{\sum_{i=1}^{N} x_{2i}}{\sqrt{N}} + \lim_{N\to\infty} \frac{\sum_{i=1}^{N} x_{2i+1}}{\sqrt{N}} &= \lim_{N\to\infty} \frac{\sum_{i=1}^{2N} x_i}{\sqrt{N}} \\
Z_{\text{even}} + Z_{\text{odd}} &= \lim_{N\to\infty} \frac{\sqrt{2}\,\sum_{i=1}^{2N} x_i}{\sqrt{2N}} \\
&= \sqrt{2}\, Z_{\text{both}}.
\end{aligned}$$

It follows that

$$P(Z_{\text{even}} + Z_{\text{odd}} \leq a) = P(\sqrt{2}\, Z_{\text{both}} \leq a).$$

Central Limit Theorem (contd.)

Given the above, and since the p.d.f. of a sum of independent r.v.'s is the convolution of the p.d.f.'s of its addends, it follows that

$$f_Z * f_Z = \frac{f_Z(g^{-1}(z))}{g'(g^{-1}(z))}$$

where $g(z) = \sqrt{2}\, z$. It follows that f_Z characterizing Z, the sum of an infinite number of samples of X, must satisfy

$$f_Z * f_Z = \frac{1}{\sqrt{2}} f_Z\left(\frac{z}{\sqrt{2}}\right).$$

Convolution of Two Gaussians

For independent r.v.'s X and Y, the distribution $Z = X + Y$ equals the convolution of f_X and f_Y:

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

Substituting the normal density

$$N(0,1) \ = \ f_X(x) = f_Y(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

into the convolution integral yields

$$f_Z(z) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{(z-x)^2}{2}} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}\, dx.$$

Convolution of Two Gaussians (contd.)

$$
\begin{aligned}
f_Z(z) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{(z-x)^2}{2}} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}\, dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2(2)}\right] \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{2\left(x-\frac{z}{2}\right)^2}{2}\right] dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi(2)}} \exp\left[-\frac{z^2}{2(2)}\right] \frac{1}{\sqrt{2\pi}\left(\frac{1}{\sqrt{2}}\right)} \exp\left[-\frac{\left(x-\frac{z}{2}\right)^2}{2\left(\frac{1}{\sqrt{2}}\right)^2}\right] dx \\
&= \frac{1}{\sqrt{2\pi(2)}} \exp\left[-\frac{z^2}{2(2)}\right] \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\left(\frac{1}{\sqrt{2}}\right)} \exp\left[-\frac{\left(x-\frac{z}{2}\right)^2}{2\left(\frac{1}{\sqrt{2}}\right)^2}\right] dx}_{1} \\
&= \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2\pi}}\, e^{-(z/\sqrt{2})^2/2} \\
&= \frac{1}{\sqrt{2}} f_X\left(\frac{z}{\sqrt{2}}\right) \\
&= N(0,2)
\end{aligned}
$$

which is normal with twice the variance.

Maximum of Exponential r.v.'s

An apple tree and a peach tree stand on a hilltop. The trees drop fruit at random times. On average, one must wait $1/p_1$ minutes for an apple to drop and $1/p_2$ minutes for a peach to drop. How many minutes on average must a person wait to collect both an apple and a peach?

$$
\begin{aligned}
\langle t_{max} \rangle &= \int_0^\infty \int_0^\infty \max(t_1,t_2) \frac{p_1 p_2}{e^{p_1 t_1} e^{p_2 t_2}} dt_1 dt_2 \\
&= \int_0^\infty \int_{t_2}^\infty t_1 \frac{p_1 p_2}{e^{p_1 t_1} e^{p_2 t_2}} dt_1 dt_2 + \int_0^\infty \int_{t_1}^\infty t_2 \frac{p_1 p_2}{e^{p_1 t_1} e^{p_2 t_2}} dt_2 dt_1 \\
&= \int_0^\infty \int_{t_2}^\infty t_1 \frac{p_1}{e^{p_1 t_1}} dt_1 \frac{p_2}{e^{p_2 t_2}} dt_2 + \int_0^\infty \int_{t_1}^\infty t_2 \frac{p_2}{e^{p_2 t_2}} dt_2 \frac{p_1}{e^{p_1 t_1}} dt_1 \\
&= \int_0^\infty \frac{p_1 t_2 + 1}{p_1 e^{p_1 t_2}} \frac{p_2}{e^{p_2 t_2}} dt_2 + \int_0^\infty \frac{p_2 t_1 + 1}{p_2 e^{p_2 t_1}} \frac{p_1}{e^{p_1 t_1}} dt_1 \\
&= \frac{1}{p_1} - \frac{p_1}{(p_1+p_2)^2} + \frac{1}{p_2} - \frac{p_2}{(p_1+p_2)^2} \\
&= \frac{1}{p_1} + \frac{1}{p_2} - \frac{1}{p_1+p_2}
\end{aligned}
$$

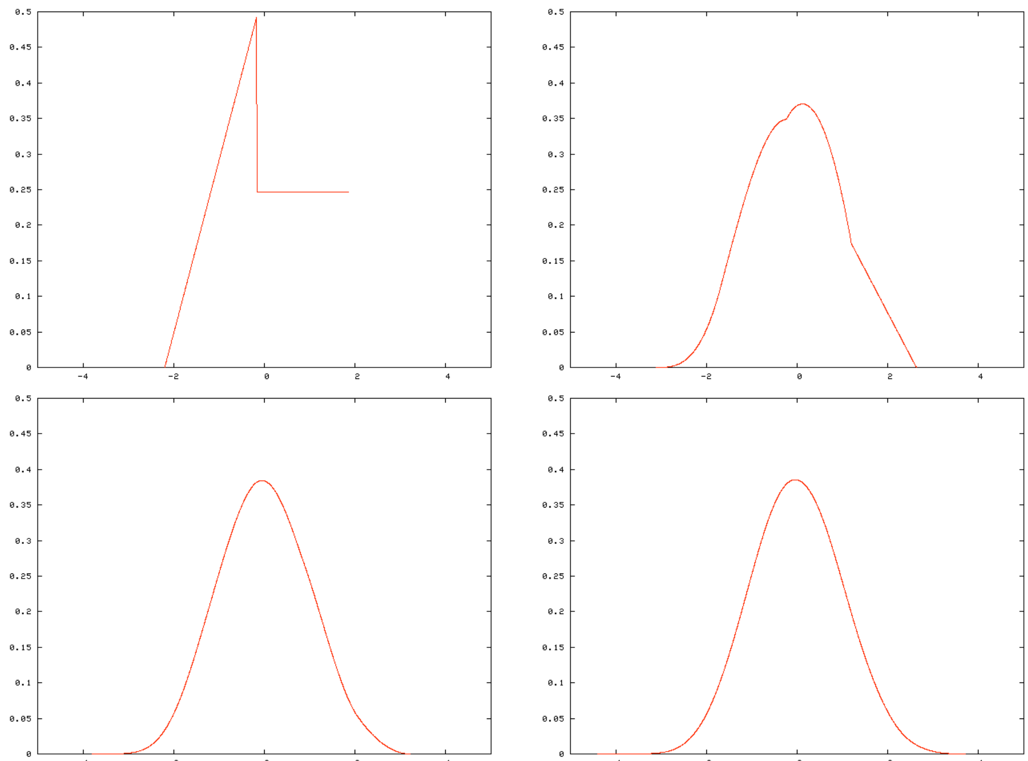

Figure 1: a) Distribution of a random variable. b) Distribution of sum of two i.i.d. random variables. c) Distribution of sum of three i.i.d. random variables. d) Distribution of sum of four i.i.d. random variables.