F. Distributions of Sums of independent random variables:

1. In general
   a. $W$ and $X$ are independent random variables
      i. Notation: $W \perp X$
   b. Produce the distribution of $Y = W + X$.
   c. The distribution of $Y$ is called the convolution of the distributions of $X$ and $Y$.
   d. Carry along $Z = X$ to make the transformation invertible.
   e. Examine the function $(y, z) = r(w, x) = (w + x, x)$.
      i. Then $w = y - z$ and $x = z$.
   f. Remove $Z$ by marginalizing.

2. The mgf of a collection of independent random variables is the product of the mgfs.
   a. $X, W$ independent, $Y = X + W$
   b. By definition, $m_Y(t) = \mathbb{E}(\exp(Yt)) = \mathbb{E}(\exp((X + W)t))$.
   c. Using properties of the exponential function,
      
   d. Using independence, $m_Y(t) = \mathbb{E}(\exp(Xt)) \mathbb{E}(\exp(Wt)) = \ldots$
\[ m_X(t)m_W(t). \]

e. One can sometimes perform convolution by recognizing the distribution from the product of the mgfs.

3. Sums of general discrete independent random variables

a. To obtain the probability function of the sum alone, marginalize by summing out \( Z \) to get

\[ p_Y(y) = \sum_{z \in \mathcal{X}} p_W(y - z)p_X(z). \]

i. Effective domain of summation may smaller than \( \mathcal{X} \), since there is no contribution if \( p_W(y - z) = 0 \).

4. Poisson Example:

a. \( X \sim \text{Pois}(\gamma), \quad W \sim \text{Pois}(\lambda), \quad X \perp W, \quad Y = X + W. \)

b. Result: \( Y \sim \text{Pois}(\lambda + \gamma) \)

c. Convolution directly:

i. probability function is

\[ p_Y(y) = \sum_{z=0}^{y} p_W(y - z)p_X(z) \]

ii. Substitute definitions:

\[ = \sum_{z=0}^{y} \exp(-\lambda(y - z))(\frac{\lambda^{y-z}}{(y-z)!}) \exp(-\gamma z)(\frac{\gamma^z}{z!}) \]

iii. Presence of \( z! \) and \( (y - z)! \) in the denominator hint that the probability might contain the sum of binomial probabilities.

- Let \( \varpi = \frac{\gamma}{\gamma + \lambda} \)
- Then
\[ p_Y(y) = \frac{\exp(-y(\lambda + \gamma))}{y!} (\gamma + \lambda)^y \]
\[ \times \sum_{z=0}^{y} \binom{y}{z} \varpi^z (1 - \varpi)^{y-z} \]
\[ = \frac{\exp(-y(\lambda + \gamma))}{y!} (\gamma + \lambda)^y. \]

d. Convolution via mgf:

i. Moment generating functions for \( X \) and \( W \) are
\[ \exp([\exp(t) - 1]\lambda) \text{ and } \exp([\exp(t) - 1]\gamma) \] resp.

ii. Product is
\[ \exp([\exp(t) - 1]\lambda) \times \exp([\exp(t) - 1]\gamma) = \exp([\exp(t) - 1][\gamma + \lambda]), \text{ the mgf for Pois(}\gamma + \lambda). \]

5. Binomial Example:

a. \( X \sim \text{Bin}(m, \pi), \ W \sim \text{Bin}(m, \pi), \ X \perp W, \ Y = X + W \)

b. Result: \( Y \sim \text{Bin}(m + m, \pi) \)

i. Require both success probabilities to be identical.

c. Convolution heuristically:

i. Construct a set of \( m + n \) successes and failures.

- Indicators are independent, probability \( \pi \)

ii. Number of \( S \) in first \( m \) trials defines \( W \).

iii. Number of \( S \) in last \( n \) trials defines \( X \).
iv. \( Y = X + W \) is the number of \( S \) over all.

v. Hence \( Y \sim \text{Bin}(m + n, \pi) \)

d. Convolution via mgf

i. \( m_X(t) = [\exp(t)\pi + 1 - \pi]^n \), \( m_W(t) = [\exp(t)\pi + 1 - \pi]^m \),
and so \( m_Y(t) = [\exp(t)\pi + 1 - \pi]^{m+n} \).

e. Convolution directly:

i. Using the general formula, \( p_Y(y) = \sum_{z=0}^{n} p_W(y - z)p_X(z) \).

ii. Substituting for the separate probability functions \( p_Y(y) = \\
\sum_{z=y-m}^{n} \binom{m}{y-z} \pi^{y-z}(1 - \pi)^{m+z-y} \binom{n}{z} \pi^{z}(1 - \pi)^{n-z} .

iii. Pull out factors not depending on \( z \), \( p_Y(y) = \\
\pi^y(1 - \pi)^{m+n-y} \sum_{z=y-m}^{n} \binom{m}{y-z} \binom{n}{z} \).

iv. The easy way to identify the sum with \( \binom{m+n}{z} \) is by already
knowing that the answer has to be binomial, either heuristically
or via the mgf.

f. Extension: If \( X_j \sim \text{Bin}(m_j, \pi) \), independent, then
\( \sum_j X_j \sim \text{Bin}(\sum_j m_j, \pi) . \)

i. \( Y_j \sim \text{Bin}(1, \pi) \), called a Bernoulli trial, independent

ii. \( X = \sum_j = 1^m Y_j \sim \text{Bin}(m, \pi) . \)

6. Negative Binomial Example:
a. \( X \sim \text{NBin}(k, \pi) \), \( W \sim \text{NBin}(m, \pi) \), \( X \perp W \), \( Y = X + W \)

b. Result: \( Y \sim \text{NBin}(m + k, \pi) \)

c. Convolution heuristically:
   i. Wait for success \( k \),
   ii. Start over and wait for success \( m \)
   iii. Has the same distribution as waiting for \( k + m \):
   \[
   Y = X + W \sim \text{NBin}(m + k, \pi).
   \]

d. Convolution directly as hard as for binomial.

e. Convolution via mgfs:
   i. \( X \) and \( W \) have mgfs \[ \exp(-t) - (1 - \pi)^{-k} \pi^k \] and \[ \exp(-t) - (1 - \pi)^{-m} \pi^m \] resp.
   ii. \( Y \) has mgf \[ \exp(-t) - (1 - \pi)^{-m+k} \pi^{k+m} \] :
   iii. \( Y \sim \text{NBin}(k + m, \pi) \).

7. A cautionary case

a. \( W \sim \sqrt{2} \text{Bin}(m, \pi) \), \( X \sim \text{Bin}(n, \pi) \), \( W \perp X \),
   \( Y = X + W \).

b. Every possible \( y \) is associated with only one possible \((x, w)\) pair, because:
   i. Suppose that \( W = w, X = x \) and \( W = u, X = v \)
• satisfy $y = x + w = u + v$.

• $w \neq u$, $x \neq v$.

ii. Let $w^*$ and $u^*$ be integers such that $w = w^*\sqrt{2}$, $u = u^*\sqrt{2}$.

iii. Then $\sqrt{2}w^* + x = \sqrt{2}u^* + v$.

iv. Then $\sqrt{2} = (v - x)/(w^* - u^*)$.

v. This is impossible, since $\sqrt{2}$ is irrational.

8. Sums of general continuous independent random variables

a. Same setup as for discrete case: $X, W$ independent,

$$Y = X + W, \ Z = X.$$ 

b. The determinant of matrix of derivatives is then

$$\begin{vmatrix} (dw, dx) \\ (dy, dz) \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix}.$$ 

c. The determinant of this matrix is then 1, and the probability density function is

$$f_{Y,Z}(y, z) = f_W(y - z)f_X(z).$$ 

d. To obtain the probability density function of the sum alone, marginalize by integrating out $Z$ to get

$$f_Y(y) = \int_X f_W(y - z)f_X(z) \, dz.$$ 

i. Effective domain of integration may smaller than $X$, since
there is no contribution if $f_W(y - z) = 0$.

ii. the integral above is called the convolution of the two probability density functions.

9. Uniform Example:

a. $X \sim \text{Unif}(0, 1)$, $W \sim \text{Unif}(0, 1)$, $X \perp W$, $Y = X + W$.

b. By general result, $f_Y(y) = \int_{[0,1]} f_W(y - z) f_X(z) \, dz$.

c. $f_X(z) = 1$ for $x \in [0, 1]$.

d. Hence $f_Y(y) = \int_{[0,1]} f_W(y - z) \, dz$.

e. The probability density function for the sum $Y$ is defined in pieces:

i. For $y \leq 1$, we know $y - z < 1$.

- $f_W(y - z) = \begin{cases} 1 & \text{if } y - z > 0 \text{ iff } y > z \\ 0 & \text{otherwise} \end{cases}$

- $f_Y(y) = \int_{[0,1]} 1 \times 1 \, dz = y$.

ii. For $y \geq 1$, we know $y - z > 0$.

- $f_W(y - z) = \begin{cases} 1 & \text{if } y - z < 1 \text{ iff } z > y - 1 \\ 0 & \text{otherwise} \end{cases}$

- $f_Y(y) = \int_{[0,1]} 1 \times 1 \, dz = 2 - y$.

f. Repeating this with more copies of $W$ gives densities for sums of arbitrary numbers of summands,

i. Becomes increasingly complex because of the need to keep all
arguments to the probability density function \( \in [0, 1] \). See Fig. 35.

*Fig. 35: Density of Convolution of 4 Uniforms*