Introduction to Probability
Sums and symmetry of random variables

Sums of RVs

If X_1, X_2, ..., X_n are n random variables, then their sum

$$\sum_{j=1}^{n} X_j = X_1 + X_2 + \cdots + X_n$$

is also a random variable.

Sums of RVs

These kinds of sums are easy to deal with if X_1, X_2, ..., X_n are

independent and identically distributed (IID):

- independent: $X_i \perp X_j$ for every $i \neq j$
- identically distributed: $X_i \sim X_j$ for every $i \neq j$, i.e. they have the same distribution

Expectation of sums of RVs

The expected value of a sum of random variables can be calculated term-by-term.

$$E\left(\sum_{j=1}^{n} X_j\right) = \sum_{j=1}^{n} E\left(X_j\right).$$

Example: Bin is sum of Bern

A $Bin(n, p)$ RV is the sum of n independent $Bern(p)$:

if $X_1, X_2, ..., X_n \sim Bern(p)$, all independent, then

$$\begin{aligned} E(X_1 + X_2 + ... + X_n) &= E(X_1) + E(X_2) + ... + E(X_n) \\ &= p + p + ... + p = np. \end{aligned}$$

Example: NegBin is sum of Geom

A NB(n,p) RV is the sum of n independent $Geom(p)$:

if $Y_1, Y_2, ..., Y_n \sim Geom(p)$, all independent, then

$$
\begin{aligned}
E(Y_1 + Y_2 + ... + Y_n) &= E(Y_1) + E(Y_2) + ... + E(Y_n) \\
&= \frac{1}{p} + \frac{1}{p} + ... + \frac{1}{p} = \frac{n}{p}.
\end{aligned}
$$

Distribution of a sum is a convolution

The distribution of a sum of random variables is determined by the **convolution** of their joint PMF (or joint PDF).

If X_1 and X_2 are discrete random variables, then the PMF of the sum $Y = X_1 + X_2$ is

$$p_Y(y) = P(Y = y) = P(X_1 + X_2 = y) = \sum_{n=-\infty}^{\infty} p_{(X_1,X_2)}(n, y - n).$$

Convolution examples: uniform

If X_1 and X_2 are the values on independent die rolls, what is the probability of their sum being 9?

Since $X_1 \perp X_2$, the convolution is easy to calculate.

$$
\begin{aligned}
P(X_1 + X_2 = 9) &= \sum_{n=-\infty}^{\infty} p_{(X_1, X_2)}(n, 9-n) \\
&= \sum_{n=3}^{6} p_{(X_1, X_2)}(n, 9-n) \\
&= \sum_{n=3}^{6} p_{X_1}(n) p_{X_2}(9-n) = \sum_{n=3}^{6} \frac{1}{36} = \frac{4}{36}.
\end{aligned}
$$

This kind of RV (discrete or continuous) is called a **triangular** RV; it is *not* uniform.

Convolution examples: binomial

Consider $X_1 \sim Bin(n, p)$ and $X_2 \sim Bin(m, p)$, with $X_1 \perp X_2$.

Then, rearranging terms and recalling the combinatorial identity*

$$\sum_{j=0}^{r} \binom{a}{j} \binom{b}{r-j} = \binom{a+b}{r},$$

the PMF of their sum is, for $k = 0, 1, 2, ..., n+m$,

$$X_1 + X_2 \sim Bin(n+m, p).$$

*and convention $\binom{a}{j} = 0$ if $j \notin \{0, ..., a\}$

Convolution examples: binomial

Proof

$$
\begin{aligned}
P(X_1 + X_2 = k) &= \sum_{j=0}^{n+m} p_{X_1}(j) p_{X_2}(k-j) \\
&= \sum_{j=0}^{n+m} \binom{n}{j} p^j (1-p)^{n-j} \binom{m}{k-j} p^{k-j} (1-p)^{m-(k-j)} \\
&= \sum_{j=0}^{n+m} \binom{n}{j} \binom{m}{k-j} p^k (1-p)^{n+m-k} \\
&= \binom{n+m}{k} p^k (1-p)^{n+m-k}. \ \blacksquare
\end{aligned}
$$

Convolution examples: Poisson

Consider $X_1 \sim \mathit{Poisson}(\lambda)$ and $X_2 \sim \mathit{Poisson}(\mu)$, with $X_1 \perp X_2$.

Then, by the binomial theorem, $X_1 + X_2 \sim \mathit{Poisson}(\lambda + \mu)$.

Proof

$$\begin{aligned}
P(X_1 + X_2 = k) &= \sum_{j=0}^{k} p_{X_1}(j) p_{X_2}(k-j) \\
&= \sum_{j=0}^{k} e^{-\lambda} \frac{\lambda^j}{j!} e^{-\mu} \frac{\mu^{k-j}}{(k-j)!} \\
&= e^{-(\lambda+\mu)} \frac{1}{k!} \sum_{j=0}^{k} \binom{k}{j} \lambda^j \mu^{k-j} = e^{-(\lambda+\mu)} \frac{(\lambda+\mu)^k}{k!}. \blacksquare
\end{aligned}$$

Convolution examples: geometric

Consider $X_1, X_2 \sim Geom(p)$ with $X_1 \perp X_2$.

Then $X_1 + X_2 \sim NB(2, p)$, a negative binomial RV.

Proof

$$
\begin{aligned}
P(X_1 + X_2 = k) &= \sum_{j=0}^{k} p_{X_1}(j) p_{X_2}(k-j) \\
&= \sum_{j=1}^{k-1} (1-p)^{j-1} p (1-p)^{k-j-1} p \\
&= p^2 \sum_{j=1}^{k-1} (1-p)^{k-2} = (k-1)(1-p)^{k-2} p^2. \blacksquare
\end{aligned}
$$

Convolution for continuous RVs

The PDF of the sum $X_1 + X_2$ of two continuous RVs X_1 and X_2 is calculated similarly, with the convolution of the joint PDF:

$$f_{X_1+X_2}(z) = \int_{-\infty}^{\infty} f_{(X_1,X_2)}(x, z-x)dx.$$

If $X_1 \perp X_2$, then their joint PDF factors, and this simplifies to

$$f_{X_1+X_2}(z) = \int_{-\infty}^{\infty} f_{X_1}(x) f_{X_2}(z-x)dx.$$

Convolution examples: normal

Consider $X_1 \sim N(\mu, \sigma^2)$ and $X_2 \sim N(\nu, \tau^2)$, with $X_1 \perp X_2$.

Then $X_1 + X_2 \sim N(\mu + \nu, \sigma^2 + \tau^2)$:

$$f_{(X_1,X_2)}(x,y) = f_{X_1}(x) f_{X_2}(y) = \frac{1}{2\pi\sigma\tau} e^{-\frac{(x-\mu)^2 + (y-\nu)^2}{2\tau^2}}$$

$$\implies f_{X_1+X_2}(z) = \frac{1}{2\pi\sigma\tau} \int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2} - \frac{((z-x)-\nu)^2}{2\tau^2}} dx$$

$=$ (lots of algebra: *completing the square*

and some calculus: *Gaussian integral*)

$$= \frac{1}{\sqrt{2\pi(\sigma^2 + \tau^2)}} e^{-\frac{(z-(\mu+\nu))^2}{2(\sigma^2+\tau^2)}}.$$

Convolution examples: exponential

Consider $X_1, X_2 \sim Exp(\lambda)$ with $X_1 \perp X_2$.

Then $X_1 + X_2 \sim Gamma(2, \lambda)$.

Proof

$$f_{(X_1,X_2)}(x,y) = f_{X_1}(x) f_{X_2}(y) = \lambda^2 e^{-\lambda(x+y)} 1_{(0,\infty)}(x) 1_{(0,\infty)}(y)$$

$$\begin{aligned} \implies f_{X_1+X_2}(z) &= \lambda^2 \int_0^z e^{-\lambda(x+(z-x))} dx \\ &= \lambda^2 e^{-\lambda z} \int_0^z dx = \lambda^2 z e^{-\lambda z} = \frac{\lambda^2}{\Gamma(2)} z^{2-1} e^{-\lambda}. \blacksquare \end{aligned}$$

Exchangeable random variables

A sequence of n random variables $(X_1, X_2, ..., X_n)$ are called **exchangeable** if, for any permutation of their indices

$$\sigma : \{1, 2, ..., n\} \to \{1, 2, ..., n\},$$

the tuple $(X_{\sigma(1)}, X_{\sigma(2)}, ..., X_{\sigma(n)})$ has the same joint PMF or PDF:

$$f_{(X_1, X_2, ..., X_n)}(x_1, x_2, ..., x_n) = f_{(X_{\sigma(1)}, X_{\sigma(2)}, ..., X_{\sigma(n)})}(x_1, x_2, ..., x_n).$$

Exchangeable random variables

For example, if X_1, X_2, X_3, X_4 are exchangeable,

then the joint PDF $f_{(X_1,X_2,X_3,X_4)}(x_1, x_2, x_3, x_4)$ is the same as:

- $f_{(X_3,X_1,X_4,X_2)}(x_1, x_2, x_3, x_4)$
- $f_{(X_4,X_2,X_3,X_1)}(x_1, x_2, x_3, x_4)$
- $f_{(X_2,X_3,X_1,X_4)}(x_1, x_2, x_3, x_4)$
- any of the other 20 joint PDF permutations.

IIDs are exchangeable; independent only is not enough

It should be clear that, if X_1, X_2, ..., X_n are IID, then they are exchangeable, since their joint PDF factors, then the factored terms can be rearranged in any order you like.

Example

If X_1, X_2, $X_3 \sim Exp(1)$ are IID, then

$$f_{X_1,X_2,X_3}(x_1, x_2, x_3) = \lambda e^{-\lambda(x_1+x_2+x_3)} = f_{X_2,X_3,X_1}(x_1, x_2, x_3).$$

IIDs are exchangeable; independent only is not enough

However, independence alone is not enough for exchangeability.

Example

If X_1, $X_2 \sim Exp(1)$ and $X_3 \sim Unif(0,4)$ are independent, then

$$f_{X_1,X_2,X_3}(x_1,x_2,x_3) = \lambda e^{-\lambda x_1} \cdot \lambda e^{-\lambda x_2} \cdot 1_{(0,4)}(x_3),$$

but

$$f_{X_2,X_3,X_1}(x_1,x_2,x_3) = \lambda e^{-\lambda x_1} \cdot 1_{(0,4)}(x_2) \cdot \lambda e^{-\lambda x_3}.$$

Hence, X_1, X_2, and X_3 are not exchangeable.

Sampling with replacement is IID

Multiple rolls of a die, flips of a coin, or draw of a card *with replacement* are all examples of IID sequences.

For example, let X_i be the value of the ith die roll in a sequence.

What is the probability that the 7th roll is a 4?

$$P(X_7 = 4) = \,?$$

IID sequences are exchangeable sequences. Hence, with no other information, the 7th roll $\sim$ the 1st roll.

$$P(X_7 = 4) = P(X_1 = 4) = \frac{1}{6}.$$

Sampling without replacement: identical, not independent

Draw twelve cards from a deck, and let X_i be the rank of card i.

Lay them on a table in order, but do not turn them over.

What is the probability that the eighth card on the table is a King?

$$P(X_8 \text{ is a } K) = \,?$$

Sampling without replacement: identical, not independent

With no evidence, each card's rank is identically distributed, and so exchangeable:

$$P(X_8 \text{ is a } K) = P(X_1 \text{ is a } K) = \frac{4}{52}.$$

With evidence, conditional probability changes.

Turning over the first card, we see the X_i are not independent.

$$P(X_8 \text{ is a } K \mid X_1 \text{ is a } K) = \frac{3}{51} \neq P(X_8 \text{ is a } K)$$
$$P(X_8 \text{ is a } K \mid X_1 \text{ is not a } K) = \frac{4}{51} \neq P(X_8 \text{ is a } K)$$

The same function of exchangeables are exchangeable

In general:

Theorem

If

$$X_1, X_2, ..., X_n$$

are exchangeable random variables, and

$$g : \mathbb{R} \to \mathbb{R}$$

is well defined, then

$$g(X_1), g(X_2), ..., g(X_n)$$

are also exchangeable.