Sums of a random number of random variables:

Setup: $X_1, X_2, \ldots$ a sequence of independent and identically distributed (i.i.d.) random variables.

N: another random variable taking values in the non-negative integers, independent of (X_i)'s.

$$S_N := X_1 + X_2 + \cdots + X_N = \sum_{i=1}^{N} X_i .$$

random number of summands.

Convention: $S_0 = 0$.

What is the distribution of S_N?

Tools: Transforms → probability generating function
→ moment " "
→ characteristic function. *

Remark: N is a version of a "stopping time".

2.

Thm: Let $X_1, X_2, \ldots$ be i.i.d. random variables, and let N be a non-negative integer-valued random variable independent of $X_1, X_2, \ldots$.

Set $S_0 := 0$ and $S_n := X_1 + X_2 + \cdots + X_n$.

Then

Composition formula

$$\varphi_{S_N}(t) = g_N(\varphi_{X_1}(t)) \quad \forall t \in \mathbb{R}$$

probability generating function of N

characteristic function of X_1. (can take any X_i because they are identically distributed)

Proof: $\varphi_{S_N}(t) \overset{\text{def}}{=} \mathbb{E}[e^{itS_N}]$

expectation with respect all X_i's and also N.

$\boxed{\text{Idea: Condition on } N}$

$= \mathbb{E}[\underbrace{\mathbb{E}[e^{itS_N} | N]}]$ law of total expectation

conditional expectation, a function of N, $h(N)$

$$= \mathbb{E}[h(N)].$$

Determine the function h.

Find the function h:

X_i's independent of N.

$$h(n) = \mathbb{E}\left[e^{itS_N} \mid N=n\right] = \mathbb{E}\left[e^{itS_n}\right]$$

$$= \mathbb{E}\left[e^{it(X_1+X_2+\dots X_n)}\right]$$

$$= \mathbb{E}\left[e^{itX_1} e^{itX_2} \cdots e^{itX_n}\right]$$

(X_i)'s are independent

$$= \underbrace{\mathbb{E}\left[e^{itX_1}\right]}_{\varphi_{X_1}(t)} \cdot \underbrace{\mathbb{E}\left[e^{itX_2}\right]}_{\varphi_{X_2}(t)} \cdots \mathbb{E}\left[e^{itX_n}\right]$$

X_i's are identically distributed.

$$= \prod_{i=1}^{n} \varphi_{X_i}(t) = \left(\varphi_{X_1}(t)\right)^n.$$

$$\Rightarrow h(n) = \left(\varphi_{X_1}(t)\right)^n \Rightarrow h(N) = \left(\varphi_{X_1}(t)\right)^N = \mathbb{E}\left[e^{itS_N} / N\right].$$

That means,

$$\varphi_{S_N}(t) = \mathbb{E}\left[\left(\varphi_{X_1}(t)\right)^N\right]$$

$$= g_N\left(\varphi_{X_1}(t)\right).$$

Recall: $g_Y(t) := \mathbb{E}\left[t^Y\right]$, Y integer-valued.

4.

Example: $X_1, X_2, \ldots$ independent and $\mathrm{Exp}(1)$ distributed.

$$f_{X_i}(x) = e^{-x}\,\mathbb{1}_{\{x>0\}}.$$

first success

$$N \in F_s(p) \qquad \mathbb{P}(N=n) = p \cdot q^{n-1}, \quad p+q=1, \; n=1,2,3,\ldots$$

Recall: $$\varphi_{X_1}(t) = \frac{1}{1-it},$$

$$g_N(t) = \frac{p \cdot t}{1-q \cdot t} \quad \text{(exercise, hint: geometric series)}$$

Thm 1

$$\Longrightarrow \varphi_{S_N}(t) = g_N(\varphi_{X_1}(t)) = \frac{p \cdot \left(\frac{1}{1-it}\right)}{1-q\left(\frac{1}{1-it}\right)}$$

$$= p\,\frac{1}{1-it-q} = \frac{p}{p-it} = \frac{1}{1-\frac{it}{p}}$$

$p+q=1$

$1-q=p$

$$= \varphi_{\mathrm{Exp}\left(\frac{1}{p}\right)}(t).$$

Answer: $S_N \in \mathrm{Exp}\left(\frac{1}{p}\right)$

Recall: $Y \in \mathrm{Exp}(a)$, $a>0$ $\Rightarrow$ $\varphi_Y(t) = \dfrac{1}{1-ita}$. Choose $a = \frac{1}{p}$.

//

Thm 2: $X_1, X_2, \dots$ iid, N independent of X_i's.

If $\mathbb{E}[N] < \infty$ and $\mathbb{E}|X_1| < \infty$

then

$$\mathbb{E}[S_N] = \mathbb{E}[N] \cdot \mathbb{E}[X_1]$$ //

Proof: First version: Use composition formula:

$$\varphi_{S_N}(t) = g_N(\varphi_{X_1}(t))$$

If the expected values exist, then

$$\mathbb{E}[S_N] = \frac{1}{i}\frac{d}{dt}\varphi_{S_N}(t)\Big|_{t=0}$$

Composition formula & chain rule

$$= \frac{1}{i}\, g_N'(\underbrace{\varphi_{X_1}(t)}_{\to 1}) \cdot \varphi_{X_1}'(t)\Big|_{t=0}$$

Recall: $\mathbb{E}[N] = g_N'(t)\Big|_{t=1}$, but $\varphi_{X_1}(t=0) = 1$.

$$= \frac{1}{\cancel{i}}\, \mathbb{E}[N] \cdot \underbrace{\varphi_{X_1}'(t=0)}_{= \cancel{i}\,\mathbb{E}[X_1]}$$

$$= \mathbb{E}[N] \cdot \mathbb{E}[X_1].$$

∎

Second proof:

$$\mathbb{E}[S_N] \overset{\text{law of total expectation}}{\overset{(*)}{=}} \mathbb{E}[\underbrace{\mathbb{E}[S_N \mid N]}_{=h(N)}]$$

$$= \mathbb{E}[h(N)]$$

with $h(n) = \mathbb{E}[S_N \mid N=n] = \mathbb{E}[S_n]$

$X_1 + X_2 + \dots + X_n$ indep of N.

$$= \mathbb{E}[X_1 + X_2 + \dots + X_n] = n \cdot \mathbb{E}[X_1]$$

X_i's are identically distributed.

Thus

$$\Rightarrow h(N) = \mathbb{E}[S_N \mid N] = N \cdot \mathbb{E}[X_1].$$

$$\overset{(*)}{\Rightarrow} \mathbb{E}[S_N] = \mathbb{E}[N \cdot \mathbb{E}[X_1]] = \mathbb{E}[N] \cdot \mathbb{E}[X_1]$$

deterministic number

Exercise: Let $X_1, X_2, \ldots,$ be iid taking values in the non-negative integers.

Let N be as above, independent from (X_i)'s.

Show that

$$g_{S_N}(t) = g_N\left(g_{X_1}(t)\right), \text{ for } |t| < 1.$$

Probability generating function of S_N.

(Solution: [AG] pages 80-81)