0.1 Probability Review - Outline

Frequency and Severity

In describing losses, we often have 3 types of random variables.

- Frequency or number of losses. This is always a discrete random variable
- Severity or amount of loss / payment. This tends to be continuous, but can be discrete, continuous, or mixed.
- Aggregate or total loss comes from summing a (Frequency) number of Severity variables. This is discrete if severity is discrete, and continuous (except at 0) if severity is continuous.

We will focus on the most common frequency distributions in this lesson.

Discrete Distributions

N is a *discrete random variable* if it has a *probability mass function* p_k,

$$p_k = \mathrm{P}[N = k]$$
$$0 \leq p_k \leq 1$$
$$\sum_k p_k = 1$$

p_k determines everything we care about. E.g.

$$\mathrm{E}[N] = \sum_k k \cdot p_k$$
$$\mathrm{E}[g(N)] = \sum_k g(k) \cdot p_k$$

For us, discrete random variables will usually be counting / frequency variables, meaning the possible values are $\{0, 1, 2, \dots\}$
Some key discrete distributions are included in tables at the exam.

Poisson

$N \sim \text{Poisson}(\lambda)$ if

$$\mathrm{P}[N = n] = e^{-\lambda} \cdot \frac{\lambda^n}{n!} \quad n = 0, 1, 2, \dots$$

Recall that the Taylor series for e^λ is

$$e^\lambda = 1 + \lambda + \frac{\lambda^2}{2!} + \cdots + \frac{\lambda^n}{n!} + \cdots$$

The $e^{-\lambda}$ term makes the probabilities to sum to 1.

$$\mathrm{E}[N] = \lambda = \mathrm{Var}[N]$$

If N_1 and N_2 are *independent* Poisson random variables then $N_1 + N_2 \sim \text{Poisson}(\lambda_1 + \lambda_2)$

Binomial

N is binomial if we have

- m independent trials
- Each trial has probability q of a loss
- $N =$ total number of losses

$$\begin{aligned} \mathrm{P}[N = n] &= \binom{m}{n} q^n (1-q)^{m-n} \\ \mathrm{E}[N] &= mq \\ \mathrm{Var}[N] &= mq(1-q) < \mathrm{E}[N] \end{aligned}$$

If N_1 and N_2 are *independent* binomial random variables with the *same* q then $N_1 + N_2 \sim \mathrm{Bin}(m_1 + m_2, q)$.

If they have different q's then the sum is gross.
If $m = 1$ we have a Bernoulli random variable

Geometric

One way to get a geometric N with mean/parameter β is if:

- Each year independently has a loss with probability $1/(1+\beta)$
- $N = \#$ of years *before* the first loss

$$\mathrm{P}[N = 0] = \frac{1}{1+\beta} \qquad \mathrm{P}[N = k] = \left(\frac{\beta}{1+\beta}\right)^k \cdot \frac{1}{1+\beta}$$

$$\begin{aligned} \mathrm{P}[N \geq k] &= \left(\frac{\beta}{1+\beta}\right)^k \\ \mathrm{E}[N] &= \beta \\ \mathrm{Var}[N] &= \beta(1+\beta) > \mathrm{E}[N] \end{aligned}$$

The geometric (like the exponential) is memoryless:

- $\mathrm{P}[N = d + n \mid N \geq d] = \mathrm{P}[N = n]$
- $\mathrm{E}[N - d \mid N \geq d] = \beta$

Negative Binomial

If $N_1, N_2, \ldots, N_r$ are iid geometric(β), then $N = \sum N_i$ is a negative binomial(r, β).
$N = \#$ years with no loss before r-th year with a loss.

$$\begin{aligned} \mathrm{P}[N=0] &= \frac{1}{(1+\beta)^r} \\ \mathrm{P}[N=k] &= \binom{r+k-1}{r-1} \cdot \left(\frac{1}{1+\beta}\right)^r \cdot \left(\frac{\beta}{1+\beta}\right)^k \\ &= \frac{r(r+1)\ldots(r+k-1)}{k!} \cdot \frac{\beta^k}{(1+\beta)^{r+k}} \\ \mathrm{E}[N] &= r\beta \quad \mathrm{Var}[N] = r\beta(1+\beta) \end{aligned}$$

These last 2 lines hold even if r is not an integer.
The coefficient in $\mathrm{P}[N=k]$ has k terms in numerator and denominator.
A geometric is a negative binomial with $r = 1$.

Comparisons

The mean and variances of these distributions come up often enough that you should memorize them.

Distribution	Mean		Variance
Binomial	mq	$>$	$mq(1-q)$
Poisson	λ	$=$	λ
Geometric	β	$<$	$\beta(1+\beta)$
Negative Binomial	$r\beta$	$<$	$r\beta(1+\beta)$

Exercise 1

The number of losses N has mean 6 and variance 24.
If given a choice between a binomial, Poisson, geometric, or negative binomial distribution, which would you choose?
What values would you use for the parameters?

Exercise 1

The number of losses N has mean 6 and variance 24.
If given a choice between a binomial, Poisson, geometric, or negative binomial distribution, which would you choose?
What values would you use for the parameters?

$\mathrm{Var}[N] > \mathrm{E}[N]$, so it is not binomial or Poisson.
For a negative binomial,

$$\begin{aligned} \mathrm{E}[N] &= r\beta = 6 \\ \mathrm{Var}[N] &= r\beta(1+\beta) = 24 \\ \frac{\mathrm{Var}[N]}{\mathrm{E}[N]} &= 1+\beta = 4 \\ \beta &= 3 \quad \Rightarrow r = 2 \end{aligned}$$

Since $r = 2$ and not 1, it is a negative binomial, not a geometric.
Alternatively, a geometric with mean 6 has variance $6(6+1) = 42$, not 24, so we knew it wasn't geometric from that as well.

Exercise 2

If N is Poisson with mean 2.4, find the mode of N.

Exercise 2

If N is Poisson with mean 2.4, find the mode of N.

We want to find what k maximizes p_k.
The data or table functions on the calculator show it is $k = 2$.
The algebraic approach is to look at p_k/p_{k-1}.
$p_k > p_{k-1}$ if and only if $p_k/p_{k-1} > 1$. The mode is the largest k such that $p_k > p_{k-1}$. But

$$\frac{p_k}{p_{k-1}} = \frac{e^{-2.4} \cdot \dfrac{2.4^k}{k!}}{e^{-2.4} \cdot \dfrac{2.4^{k-1}}{(k-1)!}} = \frac{2.4}{k}$$

$p_k/p_{k-1} > 1$ if $k < 2.4$, i.e., $k \leq 2$, and $p_k/p_{k-1} < 1$ for $k \geq 3$, making 2 the mode.
The calculator approach will be more useful on the exam.