Dr. Z.'s Probability Lecture 18 Handout:
Probability Generating Functions and The Gambler's Ruin problem

By Doron Zeilberger

Important concept: Probability generating function: If X is a **discrete** random variable taking values that are non-negative integers, then the **probability generating function** of X is

$$G_X(z) = \sum_{i=0}^{\infty} P(X=i) z^i \quad .$$

Note: Another, more abstract, way of defining $G_X(z)$ is $E[z^X]$. In that form it makes sense for *any* random variable, including continuous ones, but the most natural context is for random variables taking non-negative integer values. If we allow X to also take negative integer values, and the sample space is finite, we get a *Laurent polynomial*, where also negative powers are allowed.

Problem 18.1: If a k-faced die has faces marked with $a_1, \dots, a_k$, and the probability that it lands on a face with a_i dots is p_i, find the probability generating function of the random variable "number of dots" in the landed face.

Sol. of 18.1:

$$G_X(z) = \sum_{i=1}^{k} p_i x^{a_i} \quad .$$

Note: In particular for a **fair standard die**,

$$G_X(z) = \frac{1}{6}(z + z^2 + z^3 + z^4 + z^5 + z^6) \; = \; \frac{1}{6}\frac{z(1-z^6)}{1-z} \quad .$$

Problem 18.2: On Monday, a gambler tosses a coin whose probability of Heads is $\frac{3}{5}$.

On Tuesday,

• If Monday's outcome was Heads he rolls a four-faced die marked with the integers $\{1,2,3,4\}$ where the probability of it landing on i is $\frac{12}{25i}$ $(1 \leq i \leq 4)$.

• If Monday's outcome was Tails he rolls a four-faced die marked with the integers $\{3,4,5,6\}$ where the probability of it landing on i is $\frac{20}{19i}$ $(3 \leq i \leq 6)$.

He collects the amounts of dollars that shows on the face. Let X be his earning. Find the probability generating function of his earning. What is the probability mass function of X?

Sol. to 18.2:

The probability generating function in case he got a Heads on Monday is

$$\frac{12}{25}(z + \frac{1}{2}z^2 + \frac{1}{3}z^3 + \frac{1}{4}z^4)$$

The probability generating function in case he got a Tails on Monday is

$$\frac{20}{19}(\frac{1}{3}z^3 + \frac{1}{4}z^4 + \frac{1}{5}z^5 + \frac{1}{6}z^6) \quad .$$

Hence the probability generating function is

$$\frac{3}{5} \cdot \frac{12}{25}\left(z + \frac{1}{2}z^2 + \frac{1}{3}z^3 + \frac{1}{4}z^4\right) + \frac{2}{5} \cdot \frac{20}{19}\left(\frac{1}{3}z^3 + \frac{1}{4}z^4 + \frac{1}{5}z^5 + \frac{1}{6}z^6\right) \quad .$$

According to Maple this equals

$$\frac{36}{125}\,z + \frac{18}{125}\,z^2 + \frac{1684}{7125}\,z^3 + \frac{421}{2375}\,z^4 + \frac{8}{95}\,z^5 + \frac{4}{57}\,z^6 \quad .$$

Hence the probability mass function is

$$P(X=1) = \frac{36}{125} \quad , \quad P(X=2) = \frac{18}{125} \quad , \quad P(X=3) = \frac{1684}{7125} \quad ,$$

$$P(X=4) = \frac{421}{2375} \quad , \quad P(X=5) = \frac{8}{95} \quad , \quad P(X=6) = \frac{4}{57} \quad .$$

Of course $P(X = i) = 0$ if $i \not\in \{1, 2, 3, 4, 5, 6\}$.

Important Fact: If X and Y are any random variables with probability generating functions $G_X(z)$ and $G_Y(z)$ respectively, and X and Y are **independent**, then the probability generating function of $X + Y$, $G_{X+Y}(z)$ is their product, in other words

$$G_{X+Y}(z) = G_X(z)\,G_Y(z) \quad .$$

Note: This follows from $G_{X+Y}(z) = E[z^{X+Y}] = E[z^X \cdot z^Y] = E[z^X]\,E[z^Y]$, the latter equality following from the fact that since X and Y are independent, so are z^X and z^Y.

Similarly if you have several independent random variables, $X_1, \ldots, X_k$, we have:

$$G_{X_1+\ldots+X_k}(z) \;=\; G_{X_1}(z) \cdots G_{X_k}(z) \quad .$$

In particular, if $X_1 = X_2 = \ldots = X_k = X$ then it equals $G_X(z)^k$.

Corollaries:

1. If you toss a coin whose probability of Heads is p, n times, the probability generating function for the total number of Heads is

$$(px + (1-p))^n \quad .$$

In other words this is the probability generating function of the **Binomial Distribution** with parameters (n, p).

2. If you roll n fair standard dice n times, the probability generating function for the total number of dots is

$$\left(\frac{z(1-z^6)}{6(1-z)}\right)^n \quad .$$

Very important facts: If X is a discrete random variable with probability generating function $G_X(z)$ then

- $G_X(1) \,=\, 1$
- $E[X] \,=\, G'_X(1) \quad .$
- $E[X^2] \,=\, (z\frac{d}{dz})^2 G_X(z)|_{z=1} \quad .$

More generally

- $E[X^k] \,=\, (z\frac{d}{dz})^k G_X(z)|_{z=1} \quad .$

Problem 18.3: A drunkard walks along a straight line, with unit steps. At every moment, his probability of going left is $\frac{1}{4}$, his probability of going right is $\frac{1}{4}$, and his probability of staying in place is $\frac{1}{2}$. The decision where to go at any moment are all independent of all past or future decisions.

(i) What is the probability generating function of the random variable "location relative to starting point"?

(ii) What is the expectation?

(iii) What is the variance.

Sol. to 18.3: The probability generating function of a **single** step is

$$\frac{1}{4}z^{-1} + \frac{1}{2}z^0 + \frac{1}{4}z^1 = \frac{1}{4}z^{-1} + \frac{1}{2} + \frac{1}{4}z \quad .$$

Hence the probability generating function of his location after n step is

$$\left(\frac{1}{4}z^{-1} + \frac{1}{2} + \frac{1}{4}z\right)^n \quad .$$

Let's call it $f(z)$.

$$z\frac{d}{dz}f(z) \,=\, z(\left(\frac{1}{4}z^{-1} + \frac{1}{2} + \frac{1}{4}z\right)^n)' = zn\left(\frac{1}{4}z^{-1} + \frac{1}{2} + \frac{1}{4}z\right)^{n-1} \cdot (-\frac{1}{4}z^{-2} + \frac{1}{4})$$

$$= n\left(\frac{1}{4}z^{-1} + \frac{1}{2} + \frac{1}{4}z\right)^{n-1} \cdot (-\frac{1}{4}z^{-1} + \frac{1}{4}z) \quad .$$

In particular, $f'(1) = 0$, so $E[X] = 0$ (as it to be expected (no pun intended!), by symmetry). Next

$$(z\frac{d}{dz})^2 f(z) \; = \; n(n-1)\left(\frac{1}{4}z^{-1}+\frac{1}{2}+\frac{1}{4}z\right)^{n-2} \cdot (-\frac{1}{4}z^{-1}+\frac{1}{4}z)^2 + nz\left(\frac{1}{4}z^{-1}+\frac{1}{2}+\frac{1}{4}z\right)^{n-1} \cdot (-\frac{1}{4}z^{-1}+\frac{1}{4}z)'$$

$$= n(n-1)\left(\frac{1}{4}z^{-1}+\frac{1}{2}+\frac{1}{4}z\right)^{n-2} \cdot (-\frac{1}{4}z^{-1}+\frac{1}{4}z)^2 + n\left(\frac{1}{4}z^{-1}+\frac{1}{2}+\frac{1}{4}z\right)^{n-1} \cdot (\frac{1}{4}z^{-1}+\frac{1}{4}z) \quad .$$

Plugging-in $z = 1$ gives $E[X^2] = n/2$. Finally $Var(X) = E[X^2] - E[X]^2 = \frac{n}{2} - 0^2 = \frac{n}{2}$.

Note: A quicker way (without probability generating functions) is to find the expectation and variance of an individual step X_i, that are, respectively, 0 and $\frac{1}{4}\cdot(-1-0)^2+\frac{1}{2}\cdot(0-0)^2+\frac{1}{4}\cdot(1-0)^2 = \frac{1}{2}$, and then use the *linearity of expectation* to deduce that the expectation of the location of the drunkard after n steps is $n \cdot 0 = 0$, and use the *linearity of variance* (only valid for sums of *independent* random variables) to deduce that the variance is $n \cdot \frac{1}{2} = \frac{n}{2}$. The advantage of probability generating functions is that the same method lets you easily find higher moments, using $E[X^k] = (z\frac{d}{dz})^k f(z)^n$, (if you have Maple or whatever).

Important Classical Problem (Gambler's Ruin in A Fair Casino)

I. Probability of exiting as a winner

Suppose that right now you have $0 \le x \le N$ dollars. You play until you get broke (0 dollars) or get N dollars. At each step you win a dollar with probability $\frac{1}{2}$ and lose a dollar with probability $\frac{1}{2}$.

Q.: If right now you have x dollars, what is the probability that you exit a winner (with N dollars)?

A.: The probability is $\frac{x}{N}$.

Proof: Let's call this probability $f(x)$. Then, for $1 < x < N$

$$f(x) = \frac{1}{2}f(x-1) + \frac{1}{2}f(x+1) \quad ,$$

since if right now you have x dollars, with probability $\frac{1}{2}$ next round you would have $x-1$ dollars and with probability $\frac{1}{2}$ next round you would have $x+1$ dollars.

Of course

$$f(0) = 0 \quad , \quad f(N) = 1 \quad ,$$

since if right now you have 0 dollars, you definitely **lost** and if right now you have N dollars you definitely **won**.

This is a system of $N+1$ equations for the $N+1$ unknowns. But $g(x) = \frac{x}{N}$ also satisfies the analogous equation, since

$$g(x) = \frac{1}{2}g(x-1) + \frac{1}{2}g(x+1) \quad , \quad (check!) \quad .$$

Hence by **uniqueness** (from Linear algebra) $f(x) = g(x)$ and $f(x) = \frac{x}{N}$.

II. Expected Duration

Suppose that right now you have $0 \leq x \leq N$ dollars. You play until you get broke (0 dollars) or get N dollars. At each step you win a dollar with probability $\frac{1}{2}$ and lose a dollar with probability $\frac{1}{2}$.

Q.: If right now you have x dollars, what is the expected number of steps until you exit (either winner or loser)

A.: The expected number of steps is $x(N - x)$.

Proof: Let's call the expected number of steps $L(x)$. Then, for $1 < x < N$

$$L(x) = \frac{1}{2}L(x-1) + \frac{1}{2}L(x+1) + 1 \quad ,$$

since if right now you have x dollars, with probability $\frac{1}{2}$ next round you would have $x - 1$ dollars, and then your expected life until the end is $L(x-1)$, and with probability $\frac{1}{2}$ next round you would have $x + 1$ dollars, and then your expected life until the end is $L(x+1)$. But we have to add 1 to that, since you have performed one extra step.

Of course

$$L(0) = 0 \quad , \quad L(N) = 0 \quad .$$

since if right now you have 0 dollars, or N, the game is over (either happily or unhappily).

This is a system of $N + 1$ equations for the $N + 1$ unknowns. But $M(x) = x(N - x)$ also satisfies the analogous equation, since

$$M(x) = \frac{1}{2}M(x-1) + \frac{1}{2}M(x+1) + 1 \quad , \quad (\textit{check!}) \quad .$$

Hence by **uniqueness** (from Linear algebra) $L(x) = M(x)$ and hence $L(x) = x(N - x)$.

Problem 18.4: If you enter a fair casino with 400 dollars, and you win and lose a dollar with probability $\frac{1}{2}$, until you either get broke or have 1000 dollars. How many steps should you expect to spend there? What is the probability of exiting a winner?

Sol. to 18.4: The expectation for the number of coin-tosses is $400 \cdot (1000 - 400) = 24000$. The probability of exiting a winner is $\frac{400}{1000} = \frac{2}{5}$.

Important Classical Problem (Gambler's Ruin in an Unfair Casino)

Suppose that right now you have $0 \leq x \leq N$ dollars. You play until you get broke (0 dollars) or get N dollars. At each step you win a dollar with probability p and lose a dollar with probability $q = 1 - p$.

Q.: If right now you have x dollars, what is the probability that you exit a winner (with N dollars)?

A.: The probability is

$$\frac{1-(q/p)^x}{1-(q/p)^N} \quad ,$$

where $q = 1 - p$.

Proof: Let's call this probability $f(x)$. Then, for $1 < x < N$

$$f(x) \;=\; q\,f(x-1) + pf(x+1) \quad ,$$

since if right now you have x dollars, with probability q next round you would have $x - 1$ dollars and with probability p next round you would have $x + 1$ dollars.

Of course

$$f(0) = 0 \quad , \quad f(N) = 1 \quad ,$$

since if right now you have 0 dollars, you definitely **lost** and if right now you have N dollars you definitely **won**.

This is a system of $N + 1$ equations for the $N + 1$ unknowns. But $g(x) = \frac{1-(q/p)^x}{1-(q/p)^N}$ also satisfies the analogous equation, since

$$g(x) = q\,g(x-1) + p\,g(x+1) \quad , \quad (\textit{check}!) \quad .$$

Hence by **uniqueness** $f(x) = g(x)$ and $f(x) = \frac{1-(q/p)^x}{1-(q/p)^N}$.

Problem 18.5: If you enter a casino where your probability of winning a dollar is 0.47 and the probability of losing a dollar is 0.53, with 50 dollars, and you play until you either get 100 dollars or get broke, what is the probability that you exit a winner (with 100 dollars)?

Sol. to 18.5: Here $x = 50$, $N = 100$, $p = 0.47$ and $q = 0.53$, so the desired probability is

$$\frac{1-(0.53/0.47)^{50}}{1-(0.53/0.47)^{100}} \;=\; 0.002454889623 \quad .$$

Ans. to 18.5: The probability of exiting a winner is %0.2454889623, really tiny! So even a slight advantage for the casino will most likely **ruin** you.

Note: Suppose that the casino is fairer and your probability of winning a dollar is 0.49 and losing a dollar is 0.51, even then the probability of exiting a winner is %11.91749175.

MORAL: NEVER go to Atlantic City or Las Vegas!