Chapter 3

Probability Generating Functions

3.1 Preamble: Generating Functions

Generating functions are widely used in mathematics, and play an important role in probability theory. Consider a sequence $\{a_i : i = 0, 1, 2, ...\}$ of real numbers: the numbers can be 'parcelled up' in several kinds of 'generating functions'. The 'ordinary' *generating function* of the sequence is defined as

$$G(s) = \sum_{i=0}^{\infty} a_i s^i$$

for those values of the parameter s for which the sum converges. For a given sequence, there exists a *radius of convergence* $R(\geq 0)$ such that the sum converges absolutely if $|s| < R$ and diverges if $|s| > R$. $G(s)$ may be differentiated or integrated term by term any number of times when $|s| < R$.

For many well-defined sequences, $G(s)$ can be written in closed form, and the individual numbers in the sequence can be recovered either by series expansion or by taking derivatives.

3.2 Definitions and Properties

Consider a *count* r.v. X, i.e. a discrete r.v. taking non-negative values. Write

$$p_k = \mathrm{P}(X = k), \qquad k = 0, 1, 2, ... \tag{3.1}$$

(if X takes a finite number of values, we simply attach zero probabilities to those values which cannot occur). The *probability generating function* (PGF) of X is defined as

$$G_X(s) = \sum_{k=0}^{\infty} p_k s^k = \mathrm{E}(s^X). \tag{3.2}$$

Note that $G_X(1) = 1$, so the series converges absolutely for $|s| \leq 1$. Also $G_X(0) = p_0$.

For some of the more common distributions, the PGFs are as follows:

(i) *Constant* r.v. – if $p_c = 1$, $p_k = 0$, $k \neq c$, then

$$G_X(s) = \mathrm{E}(s^X) = s^c. \tag{3.3}$$

(ii) *Bernoulli* r.v. – if $p_1 = p$, $p_0 = 1 - p = q$, $p_k = 0$, $k \neq 0$ or 1, then

$$G_X(s) = \mathrm{E}(s^X) = q + ps. \tag{3.4}$$

(iii) *Geometric* r.v. – if $p_k = pq^{k-1}$, $k = 1, 2, ...$; $q = 1 - p$, then

$$G_X(s) = \frac{ps}{1 - qs} \quad \text{if } |s| < q^{-1} \qquad \text{(see HW Sheet 4.).} \tag{3.5}$$

(iv) *Binomial* r.v. – if $X \sim \mathrm{Bin}(n, p)$, then

$$G_X(s) = (q + ps)^n, \qquad (q = 1 - p) \qquad \text{(see HW Sheet 4.).} \tag{3.6}$$

(v) *Poisson* r.v. – if $X \sim \mathrm{Poisson}(\lambda)$, then

$$G_X(s) = \sum_{k=0}^{\infty} \frac{1}{k!} \lambda^k e^{-\lambda} s^k = e^{\lambda(s-1)}. \tag{3.7}$$

(vi) *Negative binomial* r.v. – if $X \sim \mathrm{NegBin}(n, p)$, then

$$G_X(s) = \sum_{k=0}^{\infty} \binom{k-1}{n-1} p^n q^{k-n} s^k = \left(\frac{ps}{1 - qs}\right)^n \qquad \text{if } |s| < q^{-1} \text{and } p + q = 1. \tag{3.8}$$

Uniqueness Theorem

If X and Y have PGFs G_X and G_Y respectively, then

$$\begin{array}{lrcll} & G_X(s) & = & G_Y(s) & \text{for all } s \quad (a) \\ \text{iff} & \mathrm{P}(X = k) & = & \mathrm{P}(Y = k) & \text{for } k = 0, 1, ... \quad (b) \end{array}$$

i.e. if and only if X and Y have the same probability distribution.

Proof We need only prove that (a) implies (b). The radii of convergence of G_X and G_Y are ≥ 1, so they have unique power series expansions about the origin:

$$\begin{aligned} G_X(s) &= \sum_{k=0}^{\infty} s^k \mathrm{P}(X = k) \\ G_Y(s) &= \sum_{k=0}^{\infty} s^k \mathrm{P}(Y = k). \end{aligned}$$

If $G_X = G_Y$, these two power series have identical coefficients. $\diamondsuit$

Example: If X has PGF $\dfrac{ps}{(1 - qs)}$ with $q = 1 - p$, then we can conclude that

$$X \sim \mathrm{Geometric}(p).$$

Given a function $A(s)$ which is known to be a PGF of a count r.v. X, we can obtain $p_k = \mathrm{P}(X = k)$

either by expanding $A(s)$ in a power series in s and setting

$$p_k = \text{coefficient of } s^k;$$

or by differentiating $A(s)$ k times with respect to s and setting $s = 0$.

We can extend the definition of PGF to functions of X. The PGF of $Y = H(X)$ is

$$G_Y(s) = G_{H(X)}(s) = \mathrm{E}\left(s^{H(X)}\right) = \sum_k \mathrm{P}(X = k)s^{H(k)}. \tag{3.9}$$

If H is fairly simple, it may be possible to express $G_Y(s)$ in terms of $G_X(s)$.

Example: Let $Y = a + bX$. Then

$$\begin{aligned} G_Y(s) &= \mathrm{E}(s^Y) &&= \mathrm{E}(s^{a+bX}) \\ &= s^a\mathrm{E}[(s^b)^X] &&= s^a G_X(s^b). \end{aligned} \tag{3.10}$$

3.3 Moments

Theorem

Let X be a count r.v. and $G_X^{(r)}(1)$ the rth derivative of its PGF $G_X(s)$ at $s = 1$. Then

$$\boxed{G_X^{(r)}(1) = \mathrm{E}[X(X-1)...(X-r+1)]} \tag{3.11}$$

Proof (informal)

$$\begin{aligned} G_X^{(r)}(s) &= \frac{d^r}{ds^r}[G_X(s)] \\ &= \frac{d^r}{ds^r}\left[\sum_k p_k s^k\right] \\ &= \sum_k p_k k(k-1)...(k-r+1)s^{k-r} \end{aligned}$$

(assuming the interchange of $\frac{d^r}{ds^r}$ and $\sum_k$ is justified). This series is convergent for $|s| \le 1$, so

$$G_X^{(r)}(1) = \mathrm{E}[X(X-1)...(X-r+1)], \qquad r \ge 1. \qquad \square$$

In particular:

$$G_X^{(1)}(1) \text{ (or } G'_X(1)) = \mathrm{E}(X) \tag{3.12}$$

and

$$\begin{aligned} G_X^{(2)}(1) \text{ (or } G''_X(1)) &= \mathrm{E}[X(X-1)] \\ &= \mathrm{E}(X^2) - \mathrm{E}(X) \\ &= \mathrm{Var}(X) + [\mathrm{E}(X)]^2 - \mathrm{E}(X) \end{aligned}$$

so

$$\mathrm{Var}(X) = G_X^{(2)}(1) - \left[G_X^{(1)}(1)\right]^2 + G_X^{(1)}(1). \tag{3.13}$$

Example: If $X \sim \text{Poisson}(\lambda)$, then

$$\begin{aligned} G_X(s) &= e^{\lambda(s-1)}; \\ G_X^{(1)}(s) &= \lambda e^{\lambda(s-1)} \\ \text{so} \quad & \mathrm{E}(X) = G_X^{(1)}(1) = \lambda e^0 = \lambda. \\ G_X^{(2)}(s) &= \lambda^2 e^{\lambda(s-1)} \\ \text{so} \quad & \mathrm{Var}(X) = \lambda^2 - \lambda^2 + \lambda = \lambda. \end{aligned}$$

3.4 Sums of independent random variables

Theorem

Let X and Y be independent count r.v.s with PGFs $G_X(s)$ and $G_Y(s)$ respectively, and let $Z = X + Y$. Then

$$G_Z(s) = G_{X+Y}(s) = G_X(s)G_Y(s). \tag{3.14}$$

Proof

$$\begin{aligned} G_Z(s) &= \mathrm{E}(s^Z) = \mathrm{E}(s^{X+Y}) \\ &= \mathrm{E}(s^X)\mathrm{E}(s^Y) \qquad \text{(independence)} \\ &= G_X(s)G_Y(s). \end{aligned}$$

$\diamondsuit$

Corollary

If $X_1, ..., X_n$ are independent count r.v.s with PGFs $G_{X_1}(s), ..., G_{X_n}(s)$ respectively (and n is a known integer), then

$$G_{X_1+...+X_n}(s) = G_{X_1}(s)....G_{X_n}(s). \tag{3.15}$$

Example 3.1

Find the distribution of the sum of n independent r.v.s X_i, $\quad i = 1, ...n$, where $X_i \sim \text{Poisson}(\lambda_i)$.

Solution

$$G_{X_i}(s) = e^{\lambda_i(s-1)}.$$

So

$$\begin{aligned} G_{X_1+X_2+\cdots+X_n}(s) &= \prod_{i=1}^{n} e^{\lambda_i(s-1)} \\ &= e^{(\lambda_1+\cdots+\lambda_n)(s-1)}. \end{aligned}$$

This is the PGF of a Poisson r.v., i.e.

$$\sum_{i=1}^{n} X_i \sim \text{Poisson}(\sum_{i=1}^{n} \lambda_i). \tag{3.16}$$

Example 3.2

In a sequence of n independent Bernoulli trials. Let

$$I_i = \begin{cases} 1, & \text{if the } i\text{th trial yields a success (probability } p) \\ 0, & \text{if the } i\text{th trial yields a failure (probability } q = 1 - p). \end{cases}$$

Let $X = \sum_{i=1}^{n} I_i =$ number of successes in n trials. What is the probability distribution of X?

Solution Since the trials are independent, the r.v.s $I_1, ...I_n$ are independent. So

$$G_X(s) = G_{I_1}G_{I_2}...G_{I_n}(s).$$

But $G_{I_i}(s) = q + ps$, $\quad i = 1, ...n$. So

$$G_X(s) = (q + ps)^n = \sum_{x=0}^{n} \binom{n}{x} (ps)^x q^{n-x}.$$

Then

$$\begin{aligned} \mathrm{P}(X = x) &= \text{coefficient of } s^x \text{ in } G_X(s) \\ &= \tbinom{n}{x} p^x q^{n-x}, \qquad x = 0, ..., n \end{aligned}$$

i.e. $X \sim \text{Bin}(n, p)$. $\diamondsuit$

3.5 Sum of a random number of independent r.v.s

Theorem

Let $N, X_1, X_2, ...$ be independent count r.v.s. If the $\{X_i\}$ are identically distributed, each with PGF G_X, then

$$S_N = X_1 + \cdots + X_N \qquad (3.17a)$$

has PGF

$$G_{S_N} = G_N\left(G_X(s)\right). \qquad (3.17b)$$

(*Note*: We adopt the convention that $X_1 + \cdots + X_N = 0$ for $N = 0$.) This is an example of the process known as *compounding* with respect to a parameter.

Proof We have

$$\begin{aligned}
G_{S_N}(s) &= \mathrm{E}\left(s^{S_N}\right) \\
&= \sum_{n=0}^{\infty} \mathrm{E}\left(s^{S_N}|N=n\right)\mathrm{P}(N=n) \qquad \text{(conditioning on } N) \\
&= \sum_{n=0}^{\infty} \mathrm{E}\left(s^{S_n}\right)\mathrm{P}(N=n) \\
&= \sum_{n=0}^{\infty} G_{X_1+\cdots+X_n}(s)\mathrm{P}(N=n) \\
&= \sum_{n=0}^{\infty} [G_X(s)]^n\,\mathrm{P}(N=n) \qquad \text{(using Corollary in previous section)} \\
&= G_N\left(G_X(s)\right) \qquad \text{by definition of } G_N.
\end{aligned}$$

◊

Corollaries

1. $\mathrm{E}(S_N) = \mathrm{E}(N).\mathrm{E}(X).$ (3.18)

Proof

$$\begin{aligned}
\frac{d}{ds}\left[G_{S_N}(s)\right] &= \frac{d}{ds}\left[G_N(G_X(s))\right] \\
&= \frac{dG_N(u)}{du}.\frac{du}{ds} \quad \text{where } u = G_X(s).
\end{aligned}$$

Setting $s = 1$, (so that $u = G_X(1) = 1$), we get

$$\mathrm{E}(S_N) = \left[G_N^{(1)}(1)\right].\left[G_X^{(1)}(1)\right] = \mathrm{E}(N).\mathrm{E}(X). \qquad \square$$

Similarly it can be deduced that

2. $\mathrm{Var}(S_N) = \mathrm{E}(N)\mathrm{Var}(X) + \mathrm{Var}(N)\left[\mathrm{E}(X)\right]^2.$ (3.19)

(prove!).

Example 3.3 *The Poisson Hen*

A hen lays N eggs, where $N \sim \text{Poisson}(\lambda)$. Each egg hatches with probability p, independently of the other eggs. Find the probability distribution of the number of chicks, Z.

Solution We have

$$Z = X_1 + \cdots + X_N,$$

where $X_1, \ldots$ are independent Bernoulli r.v.s with parameter p. Then

$$G_N(s) = e^{\lambda(s-1)}, \qquad G_X(s) = q + ps.$$

So

$$G_Z(s) = G_N(G_X(s)) = e^{\lambda p(s-1)}$$

– the PGF of a Poisson r.v., i.e. $Z \sim \text{Poisson}(\lambda p)$. ♢

3.6 *Using GFs to solve recurrence relations

[NOTE: Not required for examination purposes.]

Instead of appealing to the theory of difference relations when attempting to solve the recurrence relations which arise in the course of solving problems by conditioning, one can often convert the recurrence relation into a linear differential equation for a generating function, to be solved subject to appropriate boundary conditions.

Example 3.4 *The Matching Problem (yet again).*

At the end of Chapter 1 [Example (1.10)], the recurrence relation

$$p_n = \frac{n-1}{n} p_{n-1} + \frac{1}{n} p_{n-2}, \qquad n \geq 3$$

was derived for the probability p_n that no matches occur in an n-card matching problem. Solve this using an appropriate generating function.

Solution Multiply through by ns^{n-1} and sum over all suitable values of n to obtain

$$\sum_{n=3}^{\infty} ns^{n-1} p_n = s \sum_{n=3}^{\infty} (n-1)s^{n-2} p_{n-1} + s \sum_{n=3}^{\infty} s^{n-2} p_{n-2}.$$

Introducing the GF

$$G(s) = \sum_{n=1}^{\infty} p_n s^n$$

(N.B.: this is not a **P**GF, since $\{p_n : n \geq 1\}$ is not a probability distribution), we can rewrite this as

$$\begin{aligned} G'(s) - p_1 - 2p_2 s &= s[G'(s) - p_1] + sG(s) \\ \text{i.e.} \quad (1-s)G'(s) &= sG(s) + s \qquad (\text{since } p_1 = 0, p_2 = \tfrac{1}{2}). \end{aligned}$$

This first order differential equation now has to be solved subject to the boundary condition

$$G(0) = 0.$$

The result is

$$G(s) = (1-s)^{-1} e^{-s} - 1.$$

Now expand $G(s)$ as a power series in s, and extract the coefficient of s^n: this yields

$$p_n = 1 + \frac{(-1)}{1!} + \cdots \frac{(-1)^{n-1}}{(n-1)!} + \frac{(-1)^n}{n!}, \qquad n \geq 1$$

– the result obtained previously. ♢

3.7 Branching Processes

3.7.1 Definitions

Consider a hypothetical organism which lives for exactly one time unit and then dies in the process of giving birth to a family of similar organisms. We assume that:

(i) the family sizes are independent r.v.s, each taking the values 0,1,2,...;

(ii) the family sizes are identically distributed r.v.s, the number of 'children' in a family, C, having the distribution

$$\mathrm{P}(C = k) = p_k, \qquad k = 0, 1, 2, \ldots \tag{3.20}$$

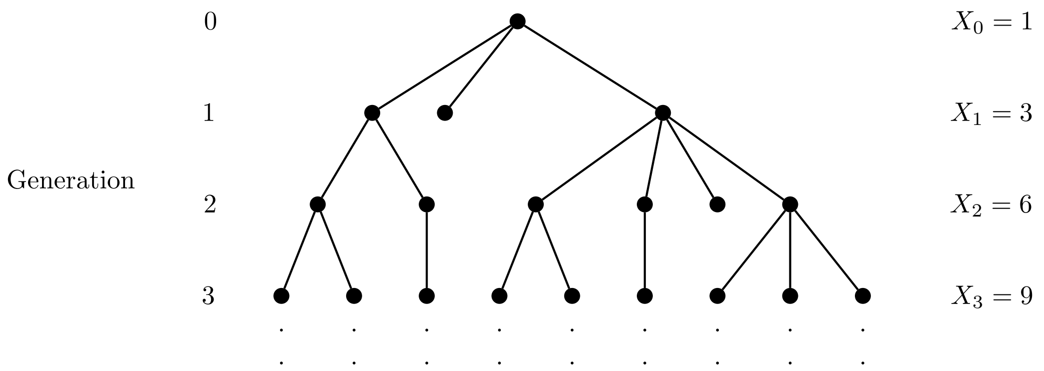
The evolution of the total population as time proceeds is termed a *branching process*: this provides a simple model for bacterial growth, spread of family names, etc.

Let

$$X_n = \text{number of organisms born at time } n \text{ (i.e., the size of the } n\text{th } generation). \tag{3.21}$$

The evolution of the population is described by the sequence of r.v.s $X_0, X_1, X_2, \ldots$ (an example of a *stochastic process* – of which more later). We assume that $X_0 = 1$, i.e. we start with precisely one organism.

A typical example can be shown as a family tree:



We can use PGFs to investigate this process.

3.7.2 Evolution of the population

Let $G(s)$ be the PGF of C:

$$G(s) = \sum_{k=0}^{\infty} \mathrm{P}(C = k)s^k, \tag{3.22}$$

and $G_n(s)$ the PGF of X_n:

$$G_n(s) = \sum_{x=0}^{\infty} \mathrm{P}(X_n = x)s^x. \tag{3.23}$$

Now $G_0(s) = s$, since $\mathrm{P}(X_0 = 1) = 1; \mathrm{P}(X_0 = x) = 0$ for $x \neq 1$, and $G_1(s) = G(s)$. Also

$$X_n = C_1 + C_2 + \cdots + C_{X_{n-1}}, \tag{3.24}$$

where C_i is the size of the family produced by the ith member of the $(n-1)$th generation.

So, since X_n is the sum of a random number of independent and identically distributed (i.i.d.) r.v.s, we have

$$G_n(s) = G_{n-1}(G(s)) \qquad \text{for } n = 2, 3, ... \tag{3.25}$$

and this is also true for $n = 1$. Iterating this result, we have

$$\begin{aligned} G_n(s) &= G_{n-1}(G(s)) \\ &= G_{n-2}(G(G(s))) \\ &= .... \\ &= G_1(G(G(....(s)....))) \\ &= G(G(G(....(s)....))), \qquad n = 0, 1, 2, ... \end{aligned}$$

i.e., G_n is the nth iterate of G.

Now

$$\begin{aligned} \mathrm{E}(X_n) &= G'_n(1) \\ &= G'_{n-1}(G(1))G'(1) \\ &= G'_{n-1}(1)G'(1) \qquad [\text{since } G(1) = 1] \end{aligned}$$

i.e.

$$\mathrm{E}(X_n) = \mathrm{E}(X_{n-1})\mu \tag{3.26}$$

where $\mu = \mathrm{E}(C)$ is the mean family size. Hence

$$\begin{aligned} \mathrm{E}(X_n) &= \mu \mathrm{E}(X_{n-1}) \\ &= \mu^2 \mathrm{E}(X_{n-2}) \\ &= .... \\ &= \mu^n \mathrm{E}(X_0) = \mu^n. \end{aligned}$$

It follows that

$$\mathrm{E}(X_n) \to \begin{cases} 0, & \text{if } \mu < 1 \\ 1, & \text{if } \mu = 1 \text{ (critical value)} \\ \infty, & \text{if } \mu > 1. \end{cases} \tag{3.27}$$

It can be shown similarly that

$$\mathrm{Var}(X_n) = \begin{cases} n\sigma^2, & \text{if } \mu = 1 \\ \sigma^2 \mu^{n-1} \dfrac{\mu^n - 1}{\mu - 1}, & \text{if } \mu \neq 1. \end{cases} \tag{3.28}$$

Example 3.5

Investigate a branching process in which C has the modified geometric distribution

$$p_k = pq^k, \qquad k = 0, 1, 2, ...; \ 0 < p = 1 - q < 1, \text{with } p \neq q \text{ [N.B.]}$$

Solution The PGF of C is

$$G(s) = \sum_{k=0}^{\infty} pq^k s^k = \frac{p}{1 - qs} \qquad \text{if } |s| < q^{-1}.$$

We now need to solve the functional equations $G_n(s) = G_{n-1}(G(s))$.

First, if $|s| \leq 1$,

$$G_1(s) = G(s) = \frac{p}{1-qs} = p\frac{(q-p)}{(q^2-p^2) - qs(q-p)}$$

(remember we have assumed that $p \neq q$). Then

$$\begin{aligned} G_2(s) = G(G(s)) &= \frac{p}{1 - \dfrac{qp}{1-qs}} \\ &= \frac{p(1-qs)}{1-qp-qs} \\ &= p\frac{(q^2-p^2) - qs(q-p)}{(q^3-p^3) - qs(q^2-p^2)}. \end{aligned}$$

We therefore conjecture that

$$G_n(s) = p\frac{(q^n-p^n) - qs(q^{n-1}-p^{n-1})}{(q^{n+1}-p^{n+1}) - qs(q^n-p^n)} \qquad \text{for } n = 1, 2, \ldots \text{ and } |s| \leq 1,$$

and this result can be proved by induction on n.

We *could* derive the entire probability distribution of X_n by expanding the r.h.s. as a power series in s, but the result is rather complicated. In particular, however,

$$\begin{aligned} \mathrm{P}(X_n = 0) = G_n(0) &= p\frac{q^n - p^n}{q^{n+1} - p^{n+1}} \\ &= \frac{\mu^n - 1}{\mu^{n+1} - 1}, \end{aligned}$$

where $\mu = q/p$ is the mean family size. It follows that ultimate extinction is certain if $\mu < 1$ and less than certain if $\mu > 1$. Separate consideration of the case $p = q = \frac{1}{2}$ (see homework) shows that ultimate extinction is also certain when $\mu = 1$. ♢

We now proceed to show that these conclusions are valid for *all* family size distributions.

3.7.3 The Probability of Extinction

The probability that the process is extinct by the nth generation is

$$e_n = \mathrm{P}(X_n = 0). \tag{3.29}$$

Now $e_n \leq 1$, and $e_n \leq e_{n+1}$ (since $X_n = 0$ implies that $X_{n+1} = 0$) - i.e. $\{e_n\}$ is a bounded monotonic sequence. So

$$e = \lim_{n\to\infty} e_n \tag{3.30}$$

exists and is called the *probability of ultimate extinction.*

Theorem 1

e is the smallest non-negative root of the equation $x = G(x)$.

Proof Note that $e_n = \mathrm{P}(X_n = 0) = G_n(0)$. Now

$$\begin{aligned} G_n(s) &= G_{n-1}(G(s)) \\ &= \ldots. \\ &= G(G....(s)....)) \\ &= G(G_{n-1}(s)). \end{aligned}$$

Set $s = 0$. Then

$$e_n = G_n(0) = G(e_{n-1}), \qquad n = 1, 2, ...$$

with boundary condition $e_0 = 0$. In the limit $n \to \infty$ we have

$$e = G(e).$$

Now let η be *any* non-negative root of the equation $x = G(x)$. G is non-decreasing on the interval $[0, 1]$ (since it has non-negative coefficients), and so

$$\begin{aligned} e_1 &= G(e_0) = G(0) \leq G(\eta) = \eta; \\ e_2 &= G(e_1) \leq G(\eta) = \eta \qquad \text{[using previous line]} \end{aligned}$$

and by induction

$$e_n \leq \eta, \qquad n = 1, 2, ...$$

So

$$e = \lim_{n\to\infty} e_n \leq \eta.$$

It follows that e is the *smallest* non-negative root. ♢

Theorem 2

$e = 1$ if and only if $\mu \leq 1$.

[*Note*: we rule out the special case $p_1 = 1; p_k = 0$ for $k \neq 1$, when $\mu = 1$ but $e = 0$]

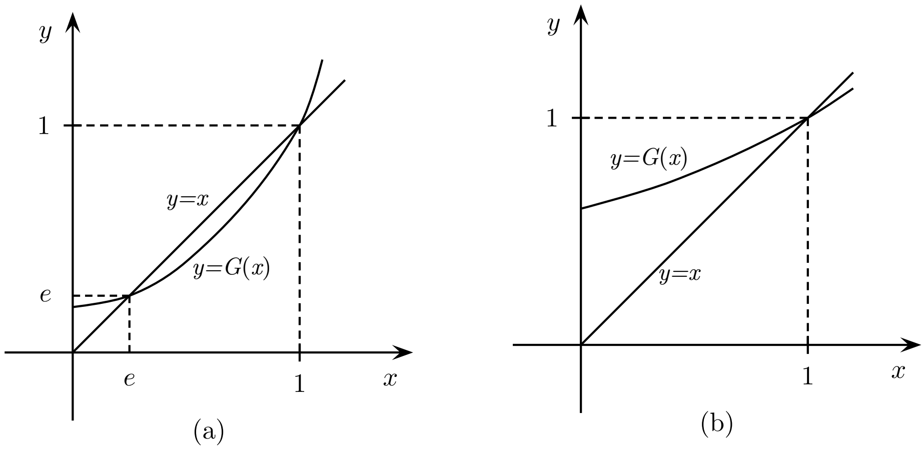
Proof We can suppose that $p_0 > 0$ (since otherwise $e = 0$ and $\mu > 1$). Now on the interval $[0, 1]$, G is

(i) continuous (since radius of convergence ≥ 1);

(ii) non-decreasing (since $G'(x) = \sum_k kp_k x^{k-1} \geq 0$);

(iii) convex (since $G''(x) = \sum_k k(k-1)p_k x^{k-2} \geq 0$).

It follows that in $[0, 1]$ the line $y = x$ has either 1 or 2 intersections with the curve $y = G(x)$,as shown below:



The curve $y = G(x)$ and the line $y = x$, in the two cases when
(a) $G'(1) > 1$ and (b) $G'(1) \leq 1$.

Since $G'(1) = \mu$, it follows that

$$e = 1 \quad \text{if and only if } \mu \leq 1.$$

□

Example 3.6

Find the probability of ultimate extinction when C has the modified geometric distribution

$$p_k = pq^k, \qquad k = 0, 1, 2, \ldots;\ 0 < p = 1 - q < 1$$

(the case $p = q$ is now permitted - cf. Example 3.5).

Solution The PGF of C is

$$G(s) = \frac{p}{1 - qs}, \qquad |s| < q^{-1},$$

and e is the smallest non-negative root of

$$G(x) = \frac{p}{1 - qx} = x.$$

The roots are

$$\frac{1 \pm (2p - 1)}{2(1 - p)}.$$

So

$$\begin{array}{lll} \text{if } p < \frac{1}{2} & (\text{i.e. } \mu = q/p > 1), & e = p/q (= \mu^{-1}); \\ \text{if } p \geq \frac{1}{2} & (\text{i.e. } \mu \leq 1), & e = 1. \end{array}$$

– in agreement with our earlier discussion in Example 3.5. ♢

[Note: in the above discussion, the continuity of probability measures (see Notes below eqn. (1.10)) has been invoked more than once without comment.]