Generating functions

The moment generating function $M_X(t)$ of a random variable X is defined via

$$M_X(t) \stackrel{\mathsf{def}}{=} \mathsf{E}\big(e^{tX}\big) \equiv \sum_{k\ge 0} \frac{\mathsf{E}X^k}{k!} t^k \,.$$

If X takes values in $\{0,1,2,\dots\}$, one also considers the (probability) generating function of X given by

$$G_X(s) \stackrel{\mathsf{def}}{=} \mathsf{E}\big(s^X\big) \equiv \sum_{n\ge 0} s^n \mathsf{P}(X=n) \,.$$

Notice that $G_X(s)$ is always finite (and analytic) for $|s|<1$, whereas $M_X(t)$ might not be defined for $t \ne 0$.

GF 1. Find the generating functions, both ordinary $G(s)$ and moment $M(t)$, for the following discrete probability distributions:

(a) the distribution describing a fair die;

(b) the distribution describing a die that always comes up 3;

(c) the uniform distribution on the set $\{n, n+1, n+2, \dots, n+k\}$;

(d) the binomial distribution on $\{n, n+1, n+2, \dots, n+k\}$;

GF 2. Let a random variable X with values in $\{0,1,2\}$ have moments $\mu_1 = 1$ and $\mu_2 = 3/2$.

(a) Find its probability generating function $G(s)$ and the probabilities $\mathsf{P}(X=k)$ for $k \in \{0,1,2\}$;

(b) find the moment generating function $M_X(t)$ and the first six moments of X.

GF 3. Let X be a discrete random variable with values in $\{0,1,2,\dots,n\}$ and moment generating function $M(t)$. Find, in terms of $M(t)$, the generating functions for:

a) $-X\,,$ b) $X+1\,,$ c) $3X\,,$ d) $aX+b\,.$

GF 4. Let X be a random variable with generating function $G_X(s)$. In terms of $G_X(s)$,

a) find the generating functions of the random variables $X+1$ and $2X$;

b) find the generating function of the sequence $\mathsf{P}(X \le n) = \sum_{k \le n} \mathsf{P}(X=k)$;

c) find the generating function of the sequence $\mathsf{P}(X=2n)$.

Hint: Which sequence has generating function $\frac{1}{2}\big(G_X(s) + G_X(-s)\big)$?

GF 5. Let the variables X and Y satisfy

$$\mathsf{P}(X=1)=\mathsf{P}(Y=5)=\frac{1}{3}\,,\qquad \mathsf{P}(X=4)=\mathsf{P}(Y=2)=\frac{2}{3}\,;$$

(a) show that X and Y have the same first and second moments, but not the same third and fourth moments;

(b) find the probability and moment generating functions for X and Y.

GF 6. Let X_1, X_2, ..., X_n be an independent trials process, with values in $\{0,1\}$ and mean $\mu=1/3$. Find the ordinary and moment generating functions for the distribution of

a) $S_1=X_1\,,$ b) $S_2=X_1+X_2\,,$ c) $S_n=X_1+\cdots+X_n\,,$

d) $A_n=S_n/n\,,$ e) $S_n^*=(S_n-n\mu)/\sqrt{n\sigma^2}\,,$

where σ^2 is the variance of X.

GF 7. A biased coin is tossed n times, each outcome is independent of all the others, and on each toss a head is shown with probability p. The total number of heads shown is X. Use the probability generating function of X to find:

(a) the mean and the variance of X;

(b) the probability that X is even;

(c) the probability that X is divisible by 3.

GF 8. A random variable X has geometric distribution with parameter p, $p\in(0,1)$, ie., $\mathsf{P}(X=k)=p(1-p)^{k-1}$, for $k=1,2,\dots$.

a) Find the generating function of X, its mean and variance.

b) Find the generating function of the sequence $\mathsf{P}(X>k)$.

GF 9. Fix $0\le\lambda\le1$, and consider two probability generating functions $G_1(s)$ and $G_2(s)$. Show that $\lambda G_1(s)+(1-\lambda)G_2(s)$ is a probability generating function, and interpret this result.

GF 10. Two magic dice are thrown independently, one showing a random number from the set $\mathcal{D}_1=\{1,3,4,5,6,8\}$ and another from $\mathcal{D}_2=\{1,2,2,3,3,4\}$.

a) Find the generating functions of the outcomes for both dice. Let T be the sum of two results; find the generating function of T.

b) Do the same for a pair of standard dice, ie., compute the generating function of the sum of two independent outcomes taken uniformly in $\mathcal{D}_0=\{1,2,3,4,5,6\}$. Compare your results.

GF 11. You have two fair tetrahedral dice whose faces are numbered 1, 2, 3, and 4. Show how to renumber the faces so that the distribution of the sum is unchanged.

GF 12. Let X_n, $n \geq 1$ be a sequence of *iid* (independent, identically distributed) random variables with values in $\{0, 1, 2, \dots\}$ and common generating function $G_X(s)$. Let $N \geq 0$ be an integer-valued random variable, independent of the sequence X_n; denote its generating function by $G_N(s)$. The sum

$$S_N \stackrel{\text{def}}{=} X_1 + X_2 + \cdots + X_N$$

of random number of X_k's has a so-called compound distribution.

a) Use the partition theorem for expectations to find $\mathsf{E}(S_N)$ in terms of $\mathsf{E}X$ and $\mathsf{E}N$;

b) Use the partition theorem for expectations to find $\mathsf{Var}(S_N)$ in terms of $\mathsf{E}X$, $\mathsf{Var}X$, $\mathsf{E}N$, and $\mathsf{Var}N$;

c) Show that the generating function $G_{S_N}(s)$ of S_N is $G_{S_N}(s) \equiv G_N\big(G_X(s)\big)$;

d) Use the previous result to compute $\mathsf{E}(S_N)$;

e) Use the result in c) to compute $\mathsf{Var}(S_N)$.

GF 13. Let $\mathbf{p}$ be the probability distribution on $\{0, 1, 2\}$ with $\mathbf{p}(1) = \frac{1}{3}$, $\mathbf{p}(2) = \frac{2}{3}$, and let $\mathbf{p}_n = \mathbf{p} \star \mathbf{p} \star \cdots \star \mathbf{p}$ be the n-fold convolution of $\mathbf{p}$ with itself.

(a) Find $\mathbf{p}_2$ by direct calculation;

(b) Find the generating functions $G(s)$ and $G_2(s)$ for $\mathbf{p}$ and $\mathbf{p}_2$, and verify that $G_2(s) = \big(G(s)\big)^2$.

(c) Find $G_n(s)$ from $G(s)$.

(d) Find the first two moments, and hence the mean and variance, of $\mathbf{p}_n$ from $G_n(s)$. Verify that the mean of $\mathbf{p}_n$ is n times the mean of $\mathbf{p}$.

(e) Find those integers j for which $\mathbf{p}_n(j) > 0$ from $G_n(s)$.

GF 14. Imagine a diligent janitor who replaces a light bulb the same day as it burns out. Suppose the first bulb is put in on day 0 and let X_i be the lifetime of the ith light bulb. For simplicity, we assume that the individual lifetimes X_i are *iid* rv's in $\{1, 2, \dots\}$ having a common distribution with generating function $G_X(s)$. In this interpretation $T_k = X_1 + X_2 + \cdots + X_k$ is the time the kth light bulb burns out; we put $T_0 = 0$. Let R_n be the event "a light bulb was replaced in n days", so that $R_0 \equiv \{T_0 = 0\} \equiv \Omega$ (a *certain event*).

a) Show that for every $n \geq 1$ we have $\mathsf{P}(R_n) = \sum_{k=1}^n \mathsf{P}\big(X_1 = k\big)\,\mathsf{P}\big(R_{n-k}\big)$;

b) Deduce that the characteristic function $G_R(s)$ of the sequence $\mathsf{P}(R_n)$ equals

$$G_R(s) = \big(1 - G_X(s)\big)^{-1}.$$

GF 15. A mature individual produces immature offspring according to the probability generating function $F(s)$.

a) Suppose we start with a population of k immature individuals, each of which grows to maturity with probability p and then reproduces, independently of the other individuals. Find the probability generating function of the number of immature individuals at the next generation.

b) Find the probability generating function of the number of mature individuals at the next generation, given that there are k mature individuals in the parent generation.

c) Show that the distributions in a) and b) above have the same mean, but not necessarily the same variance.

[**Hint:** You might prefer to first consider the case $k = 1$, and then generalise.]

GF 16. In a sequence of Bernoulli trials, let X be the number of trials up to and including ath success.

a) Show that

$$\mathsf{P}(X = r) = \binom{r-1}{a-1} p^a q^{r-a}\,, \qquad r \geq a\,.$$

b) Verify that the probability generating function for this sequence is

$$p^a s^a (1 - qs)^{-a}$$

and deduce that $\mathsf{E}X = a/p$ and $\mathsf{Var}X = aq/p^2$.

c) Show how X can be represented as a sum of a independent random variables, all with the same distribution. Use this representation to derive again the mean and the variance of X.

GF 17. Two players play a game called *heads or tails*. In this game, a coin coming up heads with probability p is tossed consecutively. Each time a head comes up Player I wins 1 pence, otherwise she loses 1 pence. Let $X_k \in \{-1, 1\}$ denote the outcome of the kth trial, and let S_n, $n \geq 0$, be the total gain of Player I after n trials, $S_n = X_1 + \cdots + X_n$, where different outcomes are assumed independent. Let T be the moment when Player I is first in the lead, ie.,

$$S_k \leq 0 \quad \text{for } k < T, \qquad \text{and} \qquad S_T = 1\,.$$

(a) find the generating function $G_T(s)$ of T;

(b) for which values of p is T a proper random variable, ie., when $\mathsf{P}(T < \infty) = 1$?

(c) compute the expectation $\mathsf{E}T$ of T;

GF 18. A slot machine operates so that at the first turn the probability for the player to win is $1/2$. Thereafter the probability for the player to win is $1/2$ if she lost at the last turn, but is $p < 1/2$ if she won at the last turn. If u_n is the probability that the player wins at the nth turn, show that for $n > 1$

$$u_n + \Bigl(\frac{1}{2} - p\Bigr)u_{n-1} = \frac{1}{2}\,.$$

Show that this equation also holds for $n = 1$, if u_0 is suitably defined, and find u_n.

GF 19. A flea randomly jumps over non-negative integers according to the following rule: a coin is flipped; if it shows "tails", the flea jumps to the next integer (ie., $k \mapsto k+1$); if the coin shows "heads", the flea jumps over the next integer (ie., $k \mapsto k+2$). Let u_n be the probability that, starting at the origin, the flea visits n at some point.

a) if the coin is fair, show that $u_n = (u_{n-1} + u_{n-2})/2$, compute the generating function $G_u(s)$ of the sequence u_n, and thus derive a formula for u_n;

b) do the same for a biased coin showing "heads" with probability $p \in [0,1]$.

GF 20. A biased coin is tossed repeatedly; on each toss, it shows a head with probability p. Let W be the number of tosses until the first occasion when r consecutive tosses have shown heads. Find the probability generating function of the random variable W.

GF 21. In a multiple-choice examination, a student chooses between one true and one false answer to each question. Assume that the student answers at random, and let N be the number of such answers until she first answers two successive questions correctly. Show that $\mathsf{E}(s^N) = s^2(4 - 2s - s^2)^{-1}$. Hence, find $\mathsf{E}(N)$ and $\mathsf{P}(N = k)$. Now find $\mathsf{E}(N)$ directly.

GF 22. In a sequence of Bernoulli trials with success probability p, let u_n be the probability that the first combination S(uccess)F(ailure) occurs (in that order) at trials number $n-1$ and n. Find the corresponding generating function, mean and variance.

GF 23. A fair coin is tossed n times. Let u_n be the probability that the sequence of tosses never has "head" followed by "head". Show that

$$u_n = \frac{1}{2}u_{n-1} + \frac{1}{4}u_{n-2}\,.$$

Find u_n, using the condition $u_0 = u_1 = 1$. Compute u_2 directly and check that your formula gives the correct value for $n = 2$.

GF 24. A biased coin is tossed N times, where N is a Poisson random variable with parameter λ. Let H be the number of heads and let T be the number of tails observed in this experiment. Use generating functions to show that H and T are Poisson random variables. Show that H and T are independent, and find the mean and the variance of $H - T$.

Optional problems [5]

GF 25.$^{\circledast}$ A loaded die may show different faces with different probabilities.

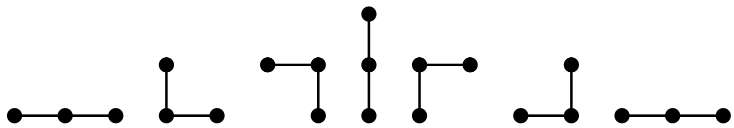
a) Is it possible to renumber two fair dice so that their sum is equally likely to take any value between 2 and 12 inclusive?

b) Can one load two traditional cubic dice in such a way that the sum of their scores is uniformly distributed on $\{2, 3, \dots, 12\}$?

c) Can the sum of three loaded dice have the same distribution as the sum of three fair dice?

GF 26.$^{\circledast}$ Successive cars pass at instants X_i seconds apart, $i \geq 1$. You require b seconds to cross the road. If the random variables X_i $(i \geq 1)$ are independent and identically distributed, find your expected waiting time until you cross. Evaluate this when $f_X(k) = qp^{k-1}$, $k \geq 1$.

GF 27.$^{\circledast}$ Let f_n be the number of paths on the integer lattice $\mathbb{Z}^2$ of length n starting from $(0, 0)$, with steps of the type $(1, 0)$, $(-1, 0)$, or $(0, 1)$, and never intersecting themselves. For instance, $f_2 = 7$, as shown in the figure below:



Show that

$$\sum_{n \geq 0} f_n s^n = \frac{1+s}{1-2s-s^2}$$

and deduce that

$$f_n = \frac{1}{2}\Big[\big(1+\sqrt{2}\big)^{n+1} + \big(1-\sqrt{2}\big)^{n+1}\Big].$$

Notice that without the restriction that the path doesn't self-intersect, there are 3^n paths with n steps. With the restriction, the number has been reduced from 3^n to roughly $(1+\sqrt{2})^n = (2.414\dots)^n$.

GF 28.$^{\circledast}$ A biased coin is tossed N times, where N is a random variable with finite mean. Show that if the numbers of heads and tails are independent, then N is Poisson.
[**Hint:** You might want to use the fact that all continuous solutions of the equation $f(x+y) = f(x)f(y)$ take the form $f(x) = e^{\lambda x}$ for some λ.]

GF 29.$^{\circledast}$ Show that for each integer n, the function

$$(s+n-1)(s+n-2)\dots s/n!$$

is the probability generating function of some random variable X. Show that $\mathsf{E}X/\log n \to 1$ as $n \to \infty$.

[5] For fun, not for exam!

GF 30.$^{\circledast}$ Let $G = (V, E)$ be a connected graph with uniformly bounded degree of its vertices (ie., for some $d \in \mathbb{N}$, every $x \in V$ is connected to at most d other vertices). Fix $x_0 \in V$, $n \in \mathbb{N}$, and consider

$$\mathcal{G}_{x_0,n} \stackrel{\text{def}}{=} \left\{ \begin{array}{c} G' = (V', E') \subset G : |V'| = n, \, x_0 \in V \\ G' - \text{connected} \end{array} \right\},$$

the collection of **all** n-vertex connected subgraphs of G containing x_0. Let $\varphi_K(s)$ be the generating function of the sequence

$$K_n = K_n(x_0) = \max \bigl|\mathcal{G}_{x_0,n}\bigr|$$

where the maximum is taken over all graphs G satisfying the conditions above.

Show that for all $s \geq 0$ we have

$$\varphi_K(s) \leq s\bigl(1 + \varphi_K(s)\bigr)^d ;$$

thus deduce that the generating function $\varphi_K(s)$ is finite for all complex s satisfying $|s| < s_0$ with $s_0 > 0$ small enough, and derive the estimate

$$K_n \leq \frac{1}{d-1} \Bigl(\frac{d^d}{(d-1)^{d-1}} \Bigr)^n .$$

GF 31.$^{\circledast}$ For $n \in \mathbb{N}$, let $g_n^{(k)}$ denote the number of k-element subsets of the set $\{1, 2, \ldots, n-1, n\}$ not containing consecutive integers. E.g., for $n = 4$ and $k = 2$ there are only three such subsets: $\{1, 3\}$, $\{1, 4\}$ and $\{2, 4\}$, so that $g_4^{(2)} = 3$. It is immediate that $g_n^{(k)} = 0$ if $n < 2k - 1$. Let $G_k(s)$ be the generating function of the sequence $g_n^{(k)}$, $G_k(s) = \sum\limits_{n=1}^{\infty} g_n^{(k)} s^n$.

(a) Show that for all n, k,

$$g_n^{(k+1)} = \sum_{j=1}^{n-2} g_{n-j-1}^{(k)} = \sum_{j=0}^{n} g_{n-j}^{(k)} \, \mathbb{1}_{\{j \geq 2\}} ,$$

ie., $g^{(k+1)}$ is the convolution of $g^{(k)}$ and $\mathbb{1}_{\{j \geq 2\}} = (0, 0, 1, 1, 1, \ldots)$; deduce that $G_k(s) = s^{2k-1}(1-s)^{-k-1}$.

(b) Expand $G_k(s)$ to get a closed expression for $g_n^{(k)}$; can you find a direct proof of your result?

(c) Use your answer in (b) to compute the probability of not seeing a pair of consecutive numbers in a (random) draw of the lottery "6 out of 49". Does it correspond to your experience? You might wish to compare your answer with the actual results, see e.g., `http://lottery.merseyworld.com/`.

GF 32.$^{\circledast}$ For $n \geq 1$ let $X_k^{(n)}$, $k = 1, \dots, n$, be Bernoulli random variables with

$$\mathsf{P}(X_k^{(n)} = 1) = 1 - \mathsf{P}(X_k^{(n)} = 0) = p_k^{(n)}.$$

Assume that $\delta^{(n)} \stackrel{\text{def}}{=} \max_{1 \leq k \leq n} p_k^{(n)} \to 0$ as $n \to \infty$ and that

$$\sum_{k=1}^{n} p_k^{(n)} \equiv \mathsf{E} \sum_{k=1}^{n} X_k^{(n)} \to \lambda \in (0, \infty).$$

Using generating functions, show that the distribution of $\sum_{k=1}^{n} X_k^{(n)}$ converges to that of a Poisson(λ) random variable. This result is known as the **law of rare events**. [**Hint:** You might wish to check that $-y^2 \leq \log(1-y) + y \leq 0$ for all $y \in [0, 1/2]$.]

GF 33.$^{\circledast}$ In n tosses of a biased coin, let L_n be the length of the longest run of heads, and set $\pi_{n,r} = \mathsf{P}(L_n < r)$. Show that

$$1 + \sum_{n=1}^{\infty} s^n \pi_{n,r} = \frac{1 - p^r s^r}{1 - s + qp^r s^{r+1}}.$$

GF 34.$^{\circledast}$ Let the generating function of the family size in an ordinary branching process be $G(s) = 1 - p(1-s)^{\beta}$, where $0 < p, \beta < 1$. Show that if $Z_0 = 1$, then

$$\mathsf{E}(s^{Z_n}) = 1 - p^{1+\beta+\cdots+\beta^{n-1}} (1-s)^{\beta^n}.$$

GF 35.$^{\circledast}$ A series of objects passes a checkpoint. Each object has (independently) probability p of being defective, and probability α of being subjected to a check, which infallibly detects a defect if it is present. Let N be the number of objects passing the checkpoint before the first defective is detected, and let D be the number of these passed objects that were defective (but undetected). Find:

(a) the joint p.g.f of D and N; (b) $\mathsf{E}(D/N)$.

If the check is not infallible, but errs with probability δ, find the above two quantities in this case.