DISCRETE DISTRIBUTIONS

Generating function (z-transform)

Definition

Let X be a discrete r.v., which take non-negative integer values, $X \in \{0, 1, 2, \ldots\}$.

Denote the point probabilities by p_i

$$p_i = \mathrm{P}\{X = i\}$$

The generating function of X denoted by $\mathcal{G}(z)$ (or $\mathcal{G}_X(z)$; also $X(z)$ or $\hat{X}(z)$) is defined by

$$\boxed{\mathcal{G}(z) = \sum_{i=0}^{\infty} p_i z^i = \mathrm{E}[z^X]}$$

Rationale:

- A handy way to record all the values $\{p_0, p_1, \ldots\}$; z is a 'bookkeeping variable'
- Often $\mathcal{G}(z)$ can be explicitly calculated (a simple analytical expression)
- When $\mathcal{G}(z)$ is given, one can conversely deduce the values $\{p_0, p_1, \ldots\}$
- Some operations on distributions correspond to much simpler operations on the generating functions
- Often simplifies the solution of recursive equations

Inverse transformation

The problem is to infer the probabilities p_i, when $\mathcal{G}(z)$ is given.

Three methods

1. Develop $\mathcal{G}(z)$ in a power series, from which the p_i can be identified as the coefficients of the z^i. The coefficients can also be calculated by derivation
$$p_i = \frac{1}{i!} \frac{d^i \mathcal{G}(z)}{dz^i}\Big|_{z=0} = \frac{1}{i!}\mathcal{G}^{(i)}(0)$$
2. By inspection: decompose $\mathcal{G}(z)$ in parts the inverse transforms of which are known; e.g. the partial fractions
3. By a (path) integral on the complex plane
$$p_i = \frac{1}{2\pi i} \oint \frac{\mathcal{G}(z)}{z^{i+1}}\, dz$$
path encircling the origin (must be chosen so that the poles of $\mathcal{G}(z)$ are outside the path)

Example 1

$$\mathcal{G}(z) = \frac{1}{1-z^2} = 1 + z^2 + z^4 + \cdots$$

$$\Rightarrow \quad p_i = \begin{cases} 1 & \text{for } i \text{ even} \\ 0 & \text{for } i \text{ odd} \end{cases}$$

Example 2

$$\mathcal{G}(z) = \frac{2}{(1-z)(2-z)} = \frac{2}{1-z} - \frac{2}{2-z} = \frac{2}{1-z} - \frac{1}{1-z/2}$$

Since $\dfrac{A}{1-az}$ corresponds to sequence $A \cdot a^i$ we deduce

$$p_i = 2 \cdot (1)^i - 1 \cdot (\tfrac{1}{2})^i = 2 - (\tfrac{1}{2})^i$$

Calculating the moments of the distribution with the aid of $\mathcal{G}(z)$

Since the p_i represent a probability distribution their sum equals 1 and

$$\mathcal{G}(1) = \mathcal{G}^{(0)}(1) = \sum_{i=1}^{\infty} p_i \cdot 1^i = 1$$

By derivation one sees

$$\begin{aligned}
\mathcal{G}^{(1)}(z) &= \tfrac{d}{dz}\mathrm{E}[z^X] \\
&= \mathrm{E}[X\,z^{X-1}] \\
\mathcal{G}^{(1)}(1) &= \mathrm{E}[X]
\end{aligned}$$

By continuing in the same way one gets

$$\boxed{\mathcal{G}^{(i)}(1) = \mathrm{E}[X(X-1)\cdots(X-i+1)] = F_i}$$

where F_i is the i^{th} <u>factorial moment</u>.

The relation between factorial moments and ordinary moments (with respect to the origin)

The factorial moments $F_i = \mathrm{E}[X(X-1)\cdots(X-i+1)]$ and ordinary moments (with resect to the origin) $M_i = \mathrm{E}[X^i]$ are related by the linear equations:

$$\boxed{\left\{\begin{array}{lcl} F_1 &=& M_1 \\ F_2 &=& M_2 - M_1 \\ F_3 &=& M_3 - 3M_2 + 2M_1 \\ \vdots \end{array}\right. \qquad \left\{\begin{array}{lcl} M_1 &=& F_1 \\ M_2 &=& F_2 + F_1 \\ M_3 &=& F_3 + 3F_2 + F_1 \\ \vdots \end{array}\right.}$$

For instance,

$$F_2 = \mathcal{G}^{(2)}(1) = \mathrm{E}[X(X-1)] = \mathrm{E}[X^2] - \mathrm{E}[X]$$

$$\Rightarrow \quad M_2 = \mathrm{E}[X^2] = F_2 + F_1 = \mathcal{G}^{(2)}(1) + \mathcal{G}^{(1)}(1)$$

$$\Rightarrow \quad \mathrm{V}[X] = M_2 - M_1^2 = \mathcal{G}^{(2)}(1) + \mathcal{G}^{(1)}(1) - (\mathcal{G}^{(1)}(1))^2 = \mathcal{G}^{(2)}(1) + \mathcal{G}^{(1)}(1)(1 - \mathcal{G}^{(1)}(1))$$

Direct calculation of the moments

The moments can also be derived from the generating function directly, without recourse to the factorial moments, as follows:

$$\begin{aligned} \frac{d}{dz}\mathcal{G}(z)\Big|_{z=1} &= \mathrm{E}[X z^{X-1}]_{z=1} &= \mathrm{E}[X] \\ \frac{d}{dz}z\,\frac{d}{dz}\mathcal{G}(z)\Big|_{z=1} &= \mathrm{E}[X^2 z^{X-1}]_{z=1} &= \mathrm{E}[X^2] \end{aligned}$$

Generally,

$$\boxed{\mathrm{E}[X^i] = \frac{d}{dz}(z\,\frac{d}{dz})^{i-1}\mathcal{G}(z)\Big|_{z=1} = (z\,\frac{d}{dz})^i\mathcal{G}(z)\Big|_{z=1}}$$

Generating function of the sum of independent random variables

Let X and Y be <u>independent</u> random variables. Then

$$\begin{aligned}
\mathcal{G}_{X+Y}(z) &= \mathrm{E}[z^{X+Y}] = \mathrm{E}[z^X z^Y] \\
&= \mathrm{E}[z^X]\mathrm{E}[z^Y] \qquad\qquad \text{independence} \\
&= \mathcal{G}_X(z)\mathcal{G}_Y(z)
\end{aligned}$$

$$\boxed{\mathcal{G}_{X+Y}(z) = \mathcal{G}_X(z)\mathcal{G}_Y(z)}$$

In terms of the original discrete distributions

$$\begin{cases} p_i = \mathrm{P}\{X = i\} \\ q_j = \mathrm{P}\{Y = j\} \end{cases}$$

the distribution of the sum is obtained by convolution $p \otimes q$

$$\mathrm{P}\{X + Y = k\} = (p \otimes q)_k = \sum_{i=0}^{k} p_i q_{k-i}$$

Thus, the generating function of a distribution obtained by convolving two distributions is the product of the generating functions of the respective original distributions.

Compound distribution and its generating function

Let Y be the sum of independent, identically distributed (*i.i.d.*) random variables X_i,

$$Y = X_1 + X_2 + \cdots X_N$$

where N is a non-negative integer-valued random variable.

Denote

$$\begin{cases} \mathcal{G}_X(z) & \text{the common generating function of the } X_i \\ \mathcal{G}_N(z) & \text{the generating function of } N \end{cases}$$

We wish to calculate $\mathcal{G}_Y(z)$

$$\begin{aligned} \mathcal{G}_Y(z) &= \mathrm{E}[z^Y] \\ &= \mathrm{E}[\mathrm{E}\left[z^Y \,|\, N\right]] \\ &= \mathrm{E}[\mathrm{E}\left[z^{X_1+\cdots X_N} \,|\, N\right]] \\ &= \mathrm{E}[\mathrm{E}\left[z^{X_1} \cdots z^{X_N} \,|\, N\right]] \\ &= \mathrm{E}[\mathcal{G}_X(z)^N] \\ &= \mathcal{G}_N(\mathcal{G}_X(z)) \end{aligned}$$

$$\boxed{\mathcal{G}_Y(z) = \mathcal{G}_N(\mathcal{G}_X(z))}$$

Bernoulli distribution $X \sim \text{Bernoulli}(p)$

A simple experiment with two possible outcomes: 'success' and 'failure'.

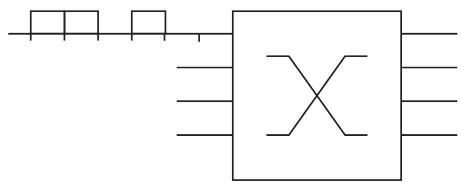
We define the random variable X as follows

$$X = \begin{cases} 1 & \text{when the experiment is successful; probability } p \\ 0 & \text{when the experiment fails; probability } q = 1 - p \end{cases}$$

Example 1. X describes the bit stream from a traffic source, which is either on or off.
The generating function

$$\begin{aligned} \mathcal{G}(z) &= p_0 z^0 + p_1 z^1 = q + pz \\ \mathrm{E}[X] &= \mathcal{G}^{(1)}(1) = p \\ \mathrm{V}[X] &= = \mathcal{G}^{(2)}(1) + \mathcal{G}^{(1)}(1 - \mathcal{G}^{(1)}) = p(1-p) = pq \end{aligned}$$

Example 2. The cell stream arriving at an input port of an ATM switch: in a time slot (cell slot) there is a cell with probability p or the slot is empty with probability q.



Binomial distribution $X \sim \mathrm{Bin}(n, p)$

X is the number of successes in a sequence of n independent Bernoulli trials.

$$X = \sum_{i=1}^{n} Y_i \qquad \text{where } Y_i \sim \text{Bernoulli}(p) \text{ and the } Y_i \text{ are independent } (i = 1, \ldots, n)$$

The generating function is obtained directly from the generating function $q + pz$ of a Bernoulli variable

$$\mathcal{G}(z) = (q + pz)^n = \sum_{i=1}^{n} \binom{n}{i} p^i (1-p)^{n-i} z^i$$

By identifying the coefficient of z^i we have

$$\boxed{p_i = \mathrm{P}\{X = i\} = \binom{n}{i} p^i (1-p)^{n-i}}$$

$$\begin{cases} \mathrm{E}[X] &= n\mathrm{E}[Y_i] = np \\ \mathrm{V}[X] &= n\mathrm{V}[Y_i] = np(1-p) \end{cases}$$

A limiting form when $\lambda = \mathrm{E}[X] = np$ is fixed and $n \to \infty$:

$$\mathcal{G}(z) = (1 - (1-z)p)^n = (1 - (1-z)\lambda/n)^n \to e^{(1-z)\lambda}$$

which is the generating function of a Poisson random variable.

The sum of binomially distributed random variables

Let the X_i $(i = 1, \ldots, k)$ be binomially distributed with the same parameter p (but with different n_i). Then the distribution of their sum is distributed as

$$X_1 + \cdots + X_k \sim \text{Bin}(n_1 + \cdots + n_k, p)$$

because the sum represents the number of successes in a sequence of $n_1 + \cdots + n_k$ identical Bernoulli trials.

Multinomial distribution

Consider a sequence of n identical trials but now each trial has k ($k \geq 2$) different outcomes. Let the probabilities of the outcomes in a single experiment be $p_1, p_2, \ldots, p_k$ ($\Sigma_{i=1}^k p_i = 1$).

Denote the number of occurrences of outcome i in the sequence by N_i. The problem is to calculate the probability $p(n_1, \ldots, n_k) = \mathrm{P}\{N_1 = n_1, \ldots, N_k = n_k\}$ of the joint event $\{N_1 = n_1, \ldots, N_k = n_k\}$.

Define the generating function of the joint distribution of several random variables $N_1, \ldots, N_k$ by

$$\mathcal{G}(z_1, \ldots, z_k) = \mathrm{E}[z_1^{N_1} \cdots z_k^{N_k}] = \sum_{n_1=0}^{\infty} \cdots \sum_{n_k=0}^{\infty} p(n_1, \ldots, n_k) z_1^{n_1} \cdots z_k^{n_k}$$

After one trial one of the N_i is 1 and the others are 0. Thus the generating function corresponding one trial is $(p_1 z_1 + \cdots + p_k z_k)$.

The generating function of n independent trials is the product of the generating functions of a single trial, i.e. $(p_1 z_1 + \cdots + p_k z_k)^n$.

From the coefficients of different powers of the z_i variables one identifies

$$\boxed{p(n_1, \ldots, n_k) = \frac{n!}{n_1! \cdots n_k!} p_1^{n_1} \cdots p_k^{n_k}} \qquad \begin{array}{l} \text{when } n_1 + \ldots + n_k = n, \\ 0 \text{ otherwise} \end{array}$$

Geometric distribution $X \sim \mathrm{Geom}(p)$

X represents the number of trials in a sequence of independent Bernoulli trials (with the probability of success p) needed until the first success occurs

$$\boxed{p_i = \mathrm{P}\{X = i\} = (1-p)^{i-1}p}$$

$i = 1, 2, \ldots$
Note that sometimes the distribution of $X - 1$ is defined to be the geometric distribution (starts from 0)

Generating function

$$\mathcal{G}(z) = p\sum_{i=1}^{\infty}(1-p)^{i-1}z^i = \frac{pz}{1-(1-p)z}$$

This can be used to calculate the expectation and the variance:

$$\begin{aligned}
\mathrm{E}[X] &= \mathcal{G}'(1) = \frac{p(1-(1-p)z) - p(1-p)z}{(1-(1-p)z)^2}\Big|_{z=1} = \frac{1}{p} \\
\mathrm{E}[X^2] &= \mathcal{G}'(1) + \mathcal{G}''(1) = \frac{1}{p} + \frac{2(1-p)}{p^2} \\
\mathrm{V}[X] &= \mathrm{E}[X^2] - \mathrm{E}[X]^2 = \frac{1-p}{p^2}
\end{aligned}$$

Geometric distribution (continued)

The probability that for the first success one needs more than n trials

$$\mathrm{P}\{X > n\} = \sum_{i=n+1}^{\infty} p_i = (1-p)^n$$

Memoryless property of geometric distribution

$$\begin{aligned}\mathrm{P}\{X > i+j \,|\, X > i\} &= \frac{\mathrm{P}\{X > i+j \cap X > i\}}{\mathrm{P}\{X > i\}} = \frac{\mathrm{P}\{X > i+j\}}{\mathrm{P}\{X > i\}} \\ &= \frac{(1-p)^{i+j}}{(1-p)^i} = \mathrm{P}\{X > j\}\end{aligned}$$

If there have been i unsuccessful trials then the probability that for the first success one needs still more than j new trials is the same as the probability that in a completely new sequence of trails one needs more than j trials for the first success.

This is as it should be, since the past trials do not have any effect on the future trials, all of which are independent.

Negative binomial distribution $X \sim \mathrm{NBin}(n, p)$

X is the number of trials needed in a sequence of Bernoulli trials needed for n successes.

If $X = i$, then among the first $(i-1)$ trials there must have been $n-1$ successes and the trial i must be a success. Thus,

$$\boxed{p_i = \mathrm{P}\{X = i\} = \binom{i-1}{n-1} p^{n-1}(1-p)^{i-n} \cdot p = \binom{i-1}{n-1} p^n (1-p)^{i-n}} \qquad \begin{array}{l} \text{if } i \geq n \\ 0 \text{ otherwise} \end{array}$$

The number of trials for the first success $\sim \mathrm{Geom}(p)$. Similarly, the number of trials needed from that point on for the next success etc. Thus,

$$X = X_1 + \cdots + X_n \qquad \text{where } X_i \sim \mathrm{Geom}(p) \quad (i.i.d.)$$

Now, the generating function of the distribution is

$$\mathcal{G}(z) = \left(\frac{pz}{1-(1-p)z}\right)^n$$

The point probabilities given above can also be deduced from this g.f.

The expectation and the variance are n times those of the geometric distribution

$$\boxed{\mathrm{E}[X] = \frac{n}{p}} \qquad \boxed{\mathrm{V}[X] = n\,\frac{1-p}{p}}$$

Poisson distribution $X \sim \mathrm{Poisson}(a)$

X is a non-negative integer-valued random variable with the point probabilities

$$\boxed{p_i = \mathrm{P}\{X = i\} = \frac{a^i}{i!}\, e^{-a}} \qquad i = 0, 1, \ldots$$

The generating function

$$\mathcal{G}(z) = \sum_{i=0}^{\infty} p_i z^i = e^{-a} \sum_{i=0}^{\infty} \frac{(za)^i}{i!} = e^{-a} e^{za}$$

$$\boxed{\mathcal{G}(z) = e^{(z-1)a}}$$

As we saw before, this generating function is obtained as a limiting form of the generating function of a $\mathrm{Bin}(n, p)$ random variable, when the average number of successes is kept fixed, $np = a$, and n tends to infinity.

Correspondingly, $X \sim \mathrm{Poisson}(\lambda t)$ represents the number of occurrences of events (e.g. arrivals) in an interval of length t from a Poisson process with intensity λ:

- the probability of an event ('success') in a small interval dt is λdt
- the probability of two simultaneous events is $\mathcal{O}(\lambda dt)$
- the number of events in disjoint intervals are independent

Poisson distribution (continued)

Poisson distribution is obeyed by e.g.

- The number of arriving calls in a given interval
- The number of calls in progress in a large (non-blocking) trunk group

Expectation and variance

$$\begin{cases} \mathrm{E}[X] = \mathcal{G}'(1) = \frac{d}{dz} e^{(z-1)a}\Big|_{z=1} = a \\ \mathrm{E}[X^2] = \mathcal{G}''(1) + \mathcal{G}'(1) = a^2 + a \end{cases} \quad \Rightarrow \quad \mathrm{V}[X] = \mathrm{E}[X^2] - \mathrm{E}[X]^2 = a^2 + a - a^2 = a$$

$$\boxed{\mathrm{E}[X] = a} \qquad \boxed{\mathrm{V}[X] = a}$$

Properties of Poisson distribution

1. The sum of Poisson random variables is Poisson distributed.

$$X = X_1 + X_2, \quad \text{where } X_1 \sim \text{Poisson}(a_1),\ X_2 \sim \text{Poisson}(a_2)$$

$$\Rightarrow \quad X \sim \text{Poisson}(a_1 + a_2)$$

Proof:

$$\mathcal{G}_{X_1}(z) = e^{(z-1)a_1},\ \mathcal{G}_{X_2}(z) = e^{(z-1)a_2}$$

$$\mathcal{G}_X(z) = \mathcal{G}_{X_1}(z)\mathcal{G}_{X_2}(z) = e^{(z-1)a_1}e^{(z-1)a_2} = e^{(z-1)(a_1+a_2)}$$

2. If the number, N, of elements in a set obeys Poisson distribution, $N \sim \text{Poisson}(a)$, and one makes a random selection with probability p (each element is independently selected with this probability), then the size of the selected set $K \sim \text{Poisson}(pa)$.

Proof: K obeys the compound distribution

$$K = X_1 + \cdots + X_N, \quad \text{where } N \sim \text{Poisson}(a) \text{ and } X_i \sim \text{Bernoulli}(p)$$

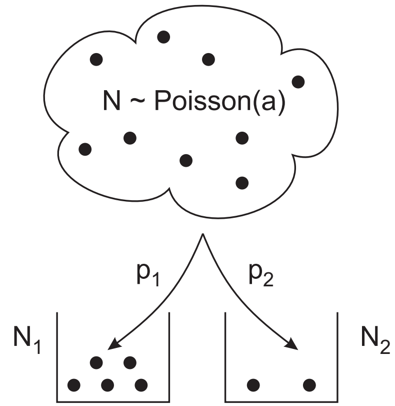
$$\mathcal{G}_X(z) = (1-p) + pz, \quad \mathcal{G}_N(z) = e^{(z-1)a}$$

$$\mathcal{G}_K(z) = \mathcal{G}_N(\mathcal{G}_X(z)) = e^{(\mathcal{G}_X(z)-1)a} = e^{[(1-p)+pz-1]a} = e^{(z-1)pa}$$

Properties of Poisson distribution (continued)

3. If the elements of a set with size $N \sim \text{Poisson}(a)$ are randomly assigned to one of two groups 1 and 2 with probabilities p_1 and $p_2 = 1-p_1$, then the sizes of the sets 1 and 2, N_1 and N_2, are independent and distributed as

   $N_1 \sim \text{Poisson}(p_1 a), \quad N_2 \sim \text{Poisson}(p_2 a)$



Proof: By the law of total probability,

$$
\begin{aligned}
\text{P}\{N_1 = n_1, N_2 = n_2\} &= \sum_{n=0}^{\infty} \underbrace{\text{P}\{N_1 = n_1, N_2 = n_2 \,|\, N = n\}}_{\text{multinomial distribution}} \quad \underbrace{\text{P}\{N = n\}}_{\text{Poisson distribution}} \\
&= \frac{n!}{n_1! n_2!} p_1^{n_1} p_2^{n_2} \cdot \frac{a^n}{n!} e^{-a}\Big|_{n=n_1+n_2} = \frac{p_1^{n_1} p_2^{n_2}}{n_1! n_2!} \cdot a^{n_1+n_2} e^{-a \overbrace{(p_1+p_2)}^{1}} \\
&= \frac{(p_1 a)^{n_1}}{n_1!} e^{-p_1 a} \cdot \frac{(p_2 a)^{n_2}}{n_2!} e^{-p_2 a} == \text{P}\{N_1 = n_1\} \cdot \text{P}\{N_2 = n_2\}
\end{aligned}
$$

The joint probability is of product form $\Rightarrow$ N_1 are N_2 independent. The factors in the product are point probabilities of Poisson$(p_1 a)$ and Poisson$(p_2 a)$ distributions.

Note, the result can be generalized for any number of sets.

Method of collective marks (Dantzig)

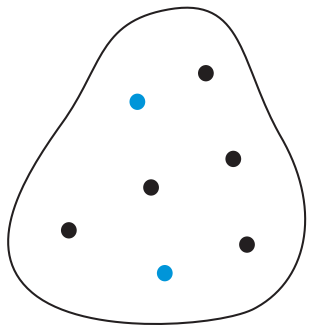
Thus far the variable z of the generating function has been considered just as a technical auxiliary variable ('book keeping variable').

In the so called method of collective marks one gives a probability interpretation for the variable z. This enables deriving some results very elegantly by simple reasoning.

Let $N = 0, 1, 2, \ldots$ be a non-negative integer-valued random variable and $\mathcal{G}_N(z)$ its generating function:

$$\mathcal{G}_N(z) = \sum_{n=0}^{\infty} p_n z^n, \qquad p_n = \mathrm{P}\{N = n\}$$

Interpretation: Think of N as representing the size of some set. Mark each of the elements in the set independently with probability $1 - z$ and leave it unmarked with probability z. *Then $\mathcal{G}_N(z)$ is the probability that there is no mark in the whole set.*

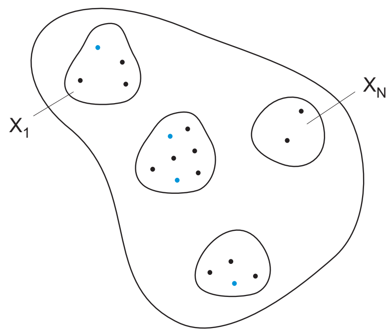


Method of collective marks (continued)

Example: The generating function of a compound distribution

$$Y = X_1 + \cdots + X_N \ , \quad \text{where}$$

$$\begin{cases} X_1 \sim X_2 \sim \cdots \sim X_N \text{ with common g.f. } \mathcal{G}_X(z) \\ N \text{ is a random variable with g.f. } \mathcal{G}_N(z) \end{cases}$$

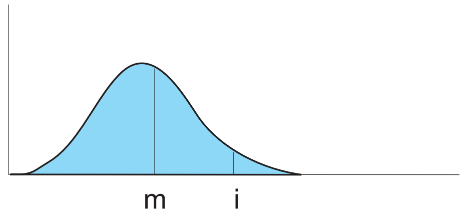


$$\begin{aligned} \mathcal{G}_Y(z) &= \mathrm{P}\{\text{none of the elements of } Y \text{ is marked}\} \\ &= \underbrace{\mathcal{G}_N(\underbrace{\mathcal{G}_X(z)}_{\substack{\text{prob. that a single} \\ \text{subset is unmarked}}})}_{\substack{\text{prob. that none of the sub-} \\ \text{sets is marked}}} \end{aligned}$$

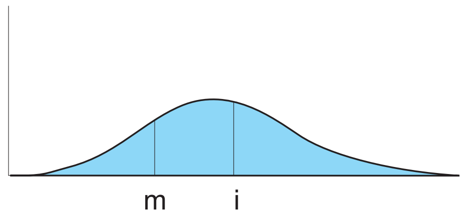
Method of probability shift: approx. calculation of point probs.

Many distributions (with large mean) can reasonably approximated by a normal distribution.

Example Poisson$(a) \approx \mathrm{N}(a, a)$, when $a \gg 1$

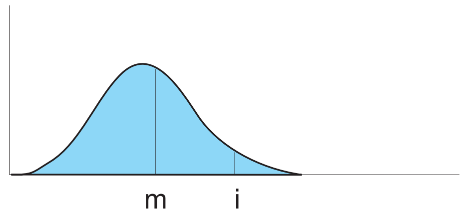


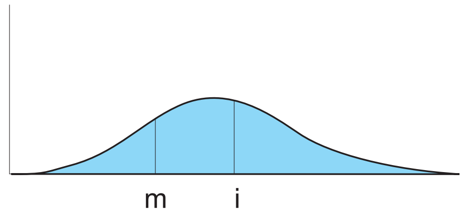
- The approximation is usually good near the mean, but far away in the tail of the distribution the relative error can be (and usually is) significant.



- The approximation can markedly be improved by the probability shift method.
- This provides a means to calculate a given point probability (in the tail) of a distribution whose generating function is known.

Probability shift (continued)





The problem is to calculate for the random variable X the point probability

$$p_i = \mathrm{P}\{X = i\}\ , \quad \text{when } i \gg \mathrm{E}[X]\ (= m)$$

In the probability shift method, one considers the (shifted) random variable X' with the point probabilities

$$p_i' = \frac{p_i z^i}{\mathcal{G}(z)}$$

These form a normed distribution, because $\mathcal{G}(z) = \sum_i p_i z^i$.

The moments of the shifted distribution are

$$\begin{cases} m'(z) = \mathrm{E}[X'] = \dfrac{1}{\mathcal{G}(z)} z \dfrac{d}{dz} \mathcal{G}(z) \\ \qquad\quad\ \mathrm{E}[X'^2] = \dfrac{1}{\mathcal{G}(z)} (z \dfrac{d}{dz})^2 \mathcal{G}(z) \\ \sigma'^2(z) = \mathrm{V}[X'] = \mathrm{E}[X'^2] - \mathrm{E}[X']^2 \end{cases}$$

Probability shift (continued)

In particular, choose the shift parameter $z = z^*$ such that $m'(z^*) = i$, i.e. so that the mean of the shifted distribution is at the point of interest i. By applying the normal approximation to the shifted distribution, one obtains

$$p'_i \approx \frac{1}{\sqrt{2\pi\sigma'^2}}$$

Conversely, by solving p_i from the previous relation one gets the desired approximation

$$\boxed{p_i \approx \frac{\mathcal{G}(z^*)}{(z^*)^i\sqrt{2\pi\sigma'^2(z^*)}}}$$

where z^* satisfies the equation $m'(z^*) = i$

In order to evaluate this expression one only needs to know the generating function of X.

The method is very useful when X is the sum of several independent random variables with different distributions, all of which (along with the corresponding generating function) are known.

The distribution of X is then complex (manyfold convolution), but as its generating function is known (the product of the respective generating functions) the above method is applicable.

Probability shift (continued)

Example (nonsensical as no approximation is really needed)

Poisson distribution

$$p_i = \frac{a^i}{i!}e^{-a}, \quad \mathcal{G}(z) = e^{(z-1)a}$$

$$p_i' = \frac{p_i z^i}{\mathcal{G}(z)} = \frac{(az)^i}{i!}e^{-az}$$ Poisson(za) distribution, so we have immediately the moments

$$\Rightarrow \quad m'(z) = az, \quad \sigma'^2(z) = az$$

The solution of the equation $m'(z^*) = i$ is $z^* = \dfrac{i}{a}$

$$p_i \approx \frac{e^{(i/a-1)a}}{(i/a)^i\sqrt{2\pi i}} = \frac{a^i}{\sqrt{2\pi i}e^{-i}i^i}e^{-a}$$

We find that the approximation gives almost exactly the correct Poisson probability but in the denominator the factorial $i!$ has been replaced by the well known Stirling approximation $i! \approx \sqrt{2\pi i}e^{-i}i^i$.