Discrete Random Variables

Dr Ahmad Khonsari
ECE Dept.
The University of Tehran

Calculating the moments of the distribution with the aid of $G(z)=\sum_{i=0}^{\infty} p_i z^i$

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

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

By derivation one sees

$$G^{(1)}(z) = \frac{d}{dz} E[z^X] = E[X z^{X-1}]$$

$$G^{(1)}(1) = E[X]$$

By continuing in the same way one gets

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

where F_i is the i^{th} factorial moment.

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:

$$
\begin{aligned}
F_1 &= M_1 & \qquad M_1 &= F_1 \\
F_2 &= M_2 - M_1 & \qquad M_2 &= F_2 + F_1 \\
F_3 &= M_3 - 3M_2 + 2M_1 & \qquad M_3 &= F_3 + 3F_2 + F_1 \\
&\vdots
\end{aligned}
$$

For instance,

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

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

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

Direct calculation of the moments

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

$$\frac{d}{dz}G(z)\bigg|_{z=1} = E[XZ^X]_{z=1} = E[X]$$

$$\frac{d}{dz}z\frac{d}{dz}G(z)\bigg|_{z=1} = E[X^2Z^X]_{z=1} = E[X^2]$$

Generally,

$$E\left[X^i\right] = \frac{d}{dz}\left(z\frac{d}{dz}\right)^{i-1} G(z)\Bigg|_{z=1} = \left(z\frac{d}{dz}\right)^{i} G(z)\Bigg|_{z=1}$$

Generating function of the sum of independent random variables

Let X and Y be independent random variables. Then

$$G_{X+Y}(z) = \underline{\mathrm{E}[z^{X+Y}]} = \mathrm{E}[z^X z^Y]$$

$$= \mathrm{E}[z^X]\mathrm{E}[z^Y] \qquad \text{independence}$$

$$= G_X(z)G_Y(z)$$

$$\boxed{G_{X+Y}(z) = G_X(z)G_Y(z)}$$

In terms of the original discrete distributions

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

$q_j = \mathrm{P}\{Y = j\}$

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

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

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 expectation

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. We wish to calculate E[Y]

$$\begin{aligned} \mathrm{E}[Y] &= \mathrm{E}[\mathrm{E}[\,Y \mid N\,]] \\ &= \mathrm{E}[N\mathrm{E}[X]] \\ &= \mathrm{E}[N]\mathrm{E}[X] \end{aligned}$$

Reminder: conditional expectation

$\mathrm{E}[\mathrm{Y}|\mathrm{X} = x] = \int_{-\infty}^{\infty} y f_{\mathrm{Y}|X}(y|x)\, d\mathrm{y}$ $\qquad$ $E[Y] = \sum_{\forall y} y p_{\mathrm{Y}|X}(\mathrm{y}|x)$

$$\mathrm{E}[Y] = \mathrm{E}\big[\mathrm{E}[\mathrm{Y}|\mathrm{X} = x]\big] = \int_{-\infty}^{\infty} \mathrm{E}[\mathrm{Y}|\mathrm{X} = x] f_X(x)\, dx$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_{\mathrm{Y}|X}(y|x)\, d\mathrm{y}\; f_X(x)\, d$$

$$= \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f_{\mathrm{X,Y}}(x,y) dx dy = \int_{-\infty}^{\infty} y f_{\mathrm{Y}}(y) dy$$

Compound distribution and its generating function

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

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

where N is a non-negative integer-valued random

variable. Denote

$G_X(z)$ the common generating function of the X_i

$G_X(z)$ the generating function of N

$$E[z^{S_N}|N = n] = [G_X(z)]^n \qquad n = 1, 2, 3, \ldots.$$

$$E[z^{S_N}|N = n] = E[z^{X_1+X_2+\cdots+X_N}|N = n]$$

$$= E[z^{X_1} z^{X_2} \ldots z^{X_N}|N = n]$$

$$= E[z^{X_1}]E[z^{X_2}] \ldots E[z^{X_n}] = [G_X(z)]^n$$

We wish to calculate $G_{S_N}(\mathrm{z})$

$$
\begin{aligned}
G_{S_N}(\mathrm{z}) &= \mathrm{E}[z^{S_N}] \\
&= \mathrm{E}[\mathrm{E}[z^{S_N}|\mathrm{N}]] && \text{Law of iterated expectation} \\
&= \mathrm{E}[G_X(z)^N] && \text{Using } \mathrm{E}[z^{S_N}|\mathrm{N}=n] = [G_X(z)]^n \\
&= G_N\left(G_X(z)\right)
\end{aligned}
$$

$$\boxed{G_{S_N}(\mathrm{z}) = G_N\left(G_X(z)\right)}$$

Analogy

$G_Y(z) = \mathrm{E}[z^Y] \quad \leftrightarrow G_N(G_X(z)) = \mathrm{E}[G_X(z)^N]$

The distribution of max and min of independent RVs

Let $X_1, X_2, \ldots, X_n$ be independent random variables (distribution functions $F_i(x)$ and tail distributions $G_i(x)$, $i = 1, \ldots, n$)

Distribution of the maximum

$P\{\max(X_1, X_2, \ldots, X_n) \le x\} = P\{X_1 \le x, \ldots, X_n \le x\}$

$= P\{X_1 \le x\} \cdots P\{X_n \le x\}$ (independence!)

$= F_1(x) \cdots F_n(x)$

Distribution of the minimum

$P\{\min(X_1, X_2, \ldots, X_n) > x\} = P\{X_1 > x, \ldots, X_n > x\}$

$= P\{X_1 > x\} \cdots P\{X_n > x\}$ (independence!)

$= G_1(x) \cdots G_n(x)$

ORDER STATISTICS

Let $X_1, X_2, \ldots, X_n$ be mutually iid continuous RVs, each having the distribution function F and density f.

Let $Y_1, Y_2, \ldots, Y_n$ be a permutation of the set $X_1, X_2, \ldots, X_n$ so as to be in increasing order.

To be specific:

$Y_1 = \min \{X_1, X_2, \ldots, X_n\}$

and

$Y_n = \max \{X_1, X_2, \ldots, X_n\}$.

Y_k is called the **k^th^-order statistic.**

Since $X_1, X_2, \ldots, X_n$ are continuous RVs, it follows that

$Y_1 < Y_2 <, \ldots < Y_n$ (as opposed to $Y_1 \leq Y_2 \leq \ldots \leq Y_n$) with a probability of one.

ORDER STATISTICS

As examples of use of order statistics, let *X;* be the lifetime of the ith component in a system of *n* independent components.

series system,

Y_1 is overall system lifetime.

Y_n is the lifetime of a parallel system

and Y_{n-m+1} is the lifetime of an m-out of-n system (the so-called N-tuple Modular Redundant or NMR system).

ORDER STATISTICS

Deriving the distribution function of Y_k

the probability that exactly

j of the X_i's lie in$(-\infty, y]$ and (n - j) lie in $(y, \infty)$is:

$\binom{n}{j} F^j(y)[l - F(y)]^{n-j}$ since the binomial distribution with parameters n and p = F(y) is applicable.

Then:

$F_{Y_k}(y) = p(y_k \leq y)$ = *P* ("at least *k* of the *X;'s* lie in the interval $(-\infty, y]$ ")

$= \sum_{j=k}^{n} \binom{n}{j} F^j(y)[l - F(y)]^{n-j}$ $\qquad -\infty \leq y \leq \infty \qquad$ (3.52)

ORDER STATISTICS

In particular, the distribution functions of Y_n and Y_1 (*i.e.* **max and min**) can be obtained from (3.52) as:

$$F_{Y_k}(y) = [F(y)]^n \qquad -\infty \leq y \leq \infty \quad ,$$

$$F_{Y_1}(y) = 1 - [1 - F(y)]^n \qquad -\infty \leq y \leq \infty$$

Thus we obtain:

$$R_{series}(t) = Ry_1(t) = 1 - F_{Y_1}(t) = [1 - F(t)]^n = [R\,(t)]^n \quad ,$$

$$R_{parallel}(t) = Ry_1(t) = 1 - F_{Y_n}(t) = 1 - [F(t)]^n = 1 - [1 - R\,(t)]^n$$

We may generalize above to the case when the lifetime distributions of individual components are distinct:

$$R_{series}(t) = Ry_1(t) = \prod_{i=1}^{n} R_{i_{Y_1}}(t)$$

7. Important distributions

We will deal with:

- discrete distributions:
  - Bernoulli;
  - binomial;
  - geometric;
  - Negative Binomial;
  - Poisson.

7.1. The Bernoulli(p) $X \sim \text{Bernoulli}(p)$

Assume we have one experiment:

event A occurs with probability p; $\Pr[\{A\}]=p$ $0 \le p \le 1$

event A does not occur with probability $(1 - p)$; $\Pr[\{\overline{A}\}]=1-p=q$ $0 \le p \le 1$

$\Omega=\{A,\overline{A}\}$

If X is a r.v. drawn from the Bernoulli(p) distribution, write: $X \sim \text{Bernoulli}(p)$ and we define RV X as:

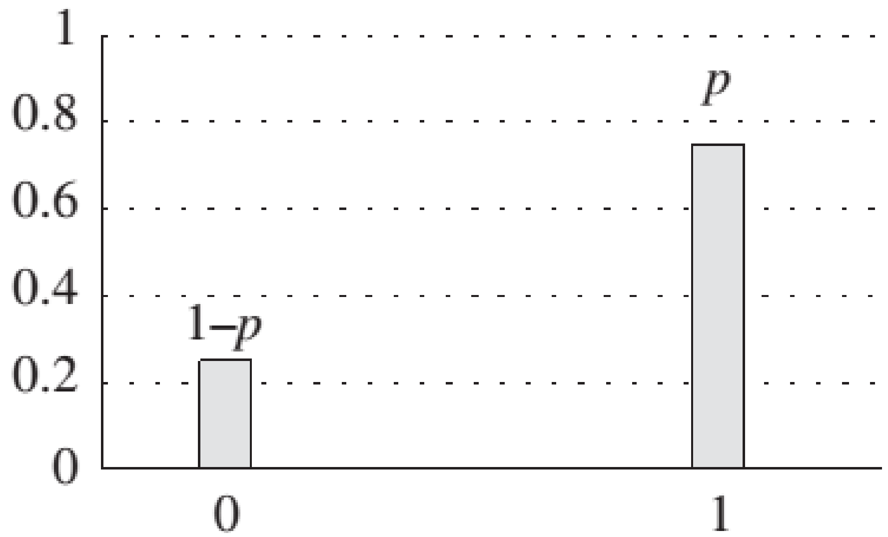
$$X = \begin{cases} 1 & \text{w/ prob } p \\ 0 & otherwise \end{cases}$$

The p.m.f. of r.v. X is defined as :

$P_X(1) = p$

$P_X(0) = 1 - p$

Example: Bernoulli(p), p=0.75



Mean and Variance of a Bernoulli Random Variable

The **mean** is:

$$\mu_X = E(X) = \sum_X xP(x) = (0)(1-P) + (1)P = P$$

And the **variance** is:

$$\sigma_X^2 = E[(X-\mu_X)^2] = \sum_X (x-\mu_X)^2 P(x)$$

$$= (0-P)^2(1-P) + (1-P)^2 P = P(1-P)$$

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

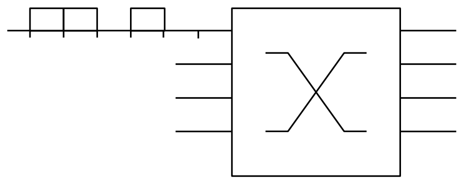
Example 1. X describes the bit stream from a traffic source, which is either on or off. The generating function

$G(z) = p_0 z^0 + p_1 z^1 = q + pz$

$E[X] = G^{(1)}(1) = p$

$V[X] = G^{(2)}(1) + G^{(1)}(1)(1 - G^{(1)}(1)) = p(1-p) = pq$

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.



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

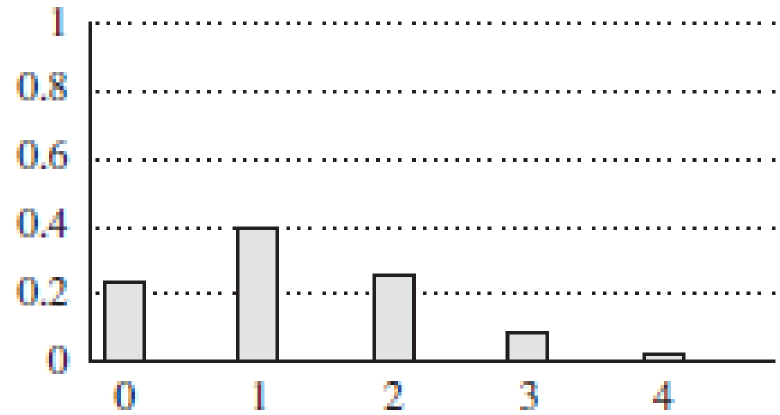
Definition: If $X \sim$ Binomial(n, p), then X represents the number of successes in n Bernoulli(p) experiments (i.e.
$X = \sum_{i=1}^{n} Y_i$ where $Y_i \sim$Bernoulli(p) and the Y_i are independent ($i = 1, . . . , n$)

The p.m.f. of r.v. X is defined as follows

$$\Pr\{X = k\} = C_n^k p^k (1-p)^{n-k} \qquad k = 0,1,\ldots,n,\ 0 \le p \le 1.$$

$$C_n^k = \binom{n}{k}$$

e.g. Probability mass function of the Binomial(n, p) distribution, with n = 4 and p = 0.3.



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

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

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

By identifying the coefficient of z^i we have

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

$$\mathrm{E}[X] = n\mathrm{E}[Y_i] = np$$

$$\mathrm{V}[X] = n\mathrm{V}[Y_i] = np(1-p)$$

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

$$G(z) = (1-(1-z)p)^n = \left(1-(1-z)\frac{\lambda}{n}\right)^n \to e^{(z-1)\lambda}$$

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

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

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 $p_1, p_2, \ldots, p_k$ ($\sum_{i=1}^{k} pi = 1$)

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

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

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

$$\boxed{p(n_1, \ldots, n_k) = \frac{n!}{n_1! \cdots n_k!} p_1^{n_1} \cdots p_k^{n_k}}$$

when $n_1 + \ldots + n_k = n$,
0 otherwise

7.3. Geometric(p) distribution

Repeating independent Bernoulli(p) experiments until the first success.
p is the probability of success

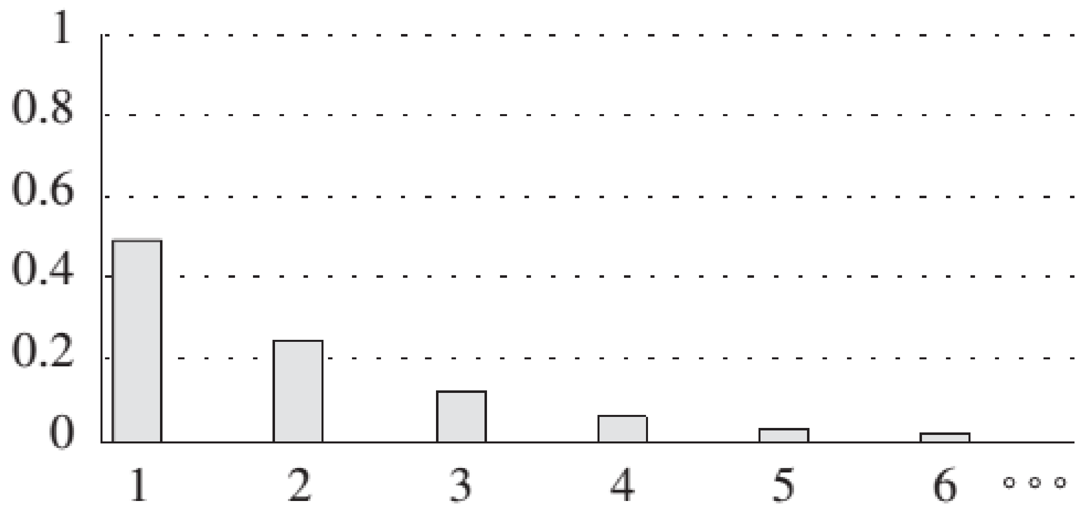
Definition: If $X \sim \text{Geometric}(p)$, then X represents the number of trials until we get a success. The p.m.f. of r.v. X is defined as follows:

$p_X(i) = \mathbf{P}\{X = i\} = (1 - p)^{i-1} p,$ where $i = 1, 2, 3, \ldots$ (105)

Mean and variance take the following form:

$$E[X] = \frac{p}{1-p} \qquad V[X] = \frac{p}{(1-p)^2} \tag{106}$$

Geometric(p), p=0.5



View 1- Shifted Geometric distribution $X \sim \text{Geom}(p)$

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

$p_i = P\{X = i\} = (1 - p)^i p \quad i = 0, 1, 2, \ldots$

$$G(z) = p \sum_{i=0}^{\infty} (1-p)^i z^i = \frac{p}{1-(1-p)z}$$

$$E[X] = \frac{1-p}{p}$$

$$V[X] = \frac{(1-p)}{p^2}$$

View 2: **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

$p_i = \mathrm{P}\{X = i\} = (1-p)^{i-1}p \qquad i = 1, 2, \ldots$

Generating function

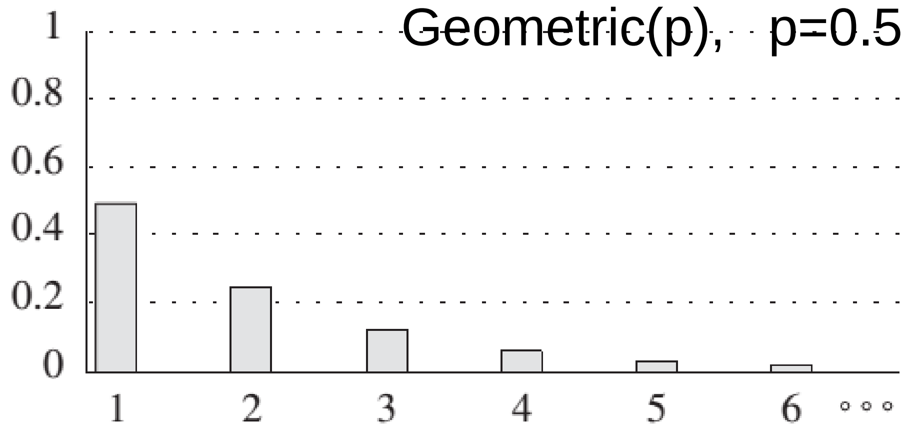
$$G(z) = p\sum_{i=1}^{\infty}(1-p)^{i-1}z^i = \frac{pz}{1-(1-p)z}$$

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

$$E[X] = G'(1) = \left.\frac{p(1-(1-p)z)+p(1-p)z}{(1-(1-p)z)^2}\right|_{z=1} = \frac{1}{p}$$

$$E[X^2] = G''(1) + G'(1) = \frac{1}{p} + \frac{2(1-p)}{p^2}$$

$$V[X] = E[X^2] - E[X]^2 = \frac{(1-p)}{p^2}$$



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

$$\mathrm{P}\{X > i+j \mid X > i\} = \frac{\mathrm{P}\{X > i+j \cap X > i\}}{\mathrm{P}\{X > i\}} = \frac{\mathrm{P}\{X > i+j\}}{\mathrm{P}\{X > i\}}$$

$$= \frac{(1-p)^{i+j}}{(1-p)^i} = \mathrm{P}\{X > j\}$$

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,

$$\Pr\{X = i\} = \binom{i-1}{n-1} p^{n-1}(1-p)^{i-n}.p = \binom{i-1}{n-1} p^{n}(1-p)^{i-n}$$

if $i \geq n$
0 otherwise

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,

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

Now, the generating function of the distribution is

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

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

$$\mathrm{E}[X] = \frac{n}{p} \qquad V[X] = n\frac{1-p}{p^2}$$

7.4. Poisson(λ) distribution

$X \sim \text{Poisson}(\lambda)$,

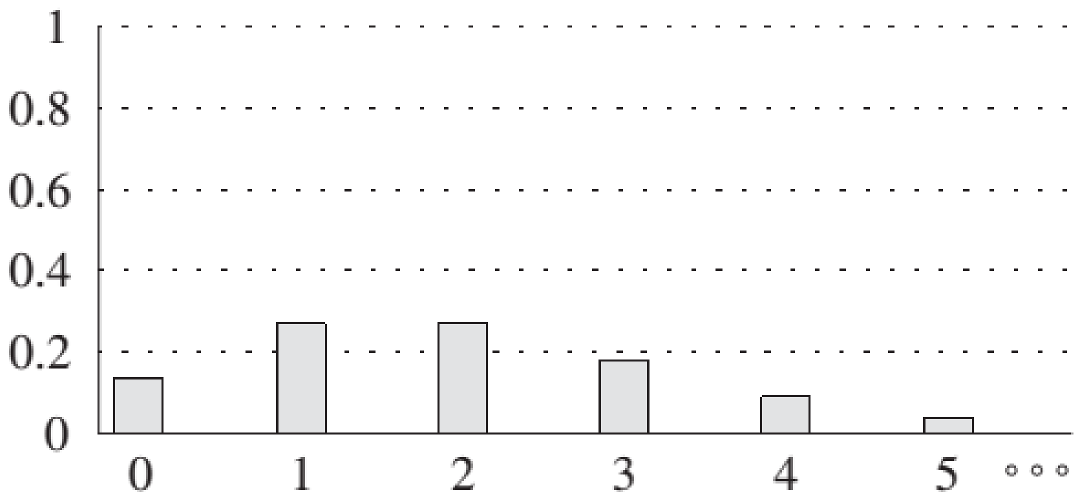
Definition: X is a non-negative integer-valued random variable with the point probabilities

$$p_i = P\{X = i\} = \frac{\lambda^i}{i!} e^{-\lambda} \quad i = 0,1, \quad (107)$$

Mean and variance are as follows:

$$E[X] = \lambda, \qquad V[X] = \lambda. \qquad (108)$$

Poisson(λ), λ =2



Poisson distribution X ~Poisson(λ)

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

The generating function

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

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

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

Correspondingly, X ~ Poisson(λt) 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 $O(\lambda dt)$
- 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

$$E[X] = G'(1) = \frac{d}{dz} e^{(z-1)\lambda}\bigg|_{z=1} = \lambda$$

$$E[X^2] = G''(1) + G'(1) = \lambda^2 + \lambda$$

$$\Rightarrow \quad V[x] = E[X^2] - E[X]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

$$E[x] = \lambda \qquad V[x] = \lambda$$

Properties of Poisson distribution

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

$$X = X_1 + X_2, \quad \text{where } X_1 \sim \text{Poisson}(\lambda_1),\ X_2 \sim \text{Poisson}(\lambda_2)$$

$$\Rightarrow \quad X \sim \text{Poisson}(\lambda_1 + \lambda_2)$$

Proof:

$$G_{X_1}(z) = e^{(z-1)\lambda_1} \quad , G_{X_2}(z) = e^{(z-1)\lambda_2}$$

$$G_X(z) = G_{X_1}(z) G_{X_2}(z) = e^{(z-1)\lambda_1} e^{(z-1)\lambda_2} = e^{(z-1)(\lambda_1 + \lambda_2)}$$

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

Proof: K obeys the compound distribution

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

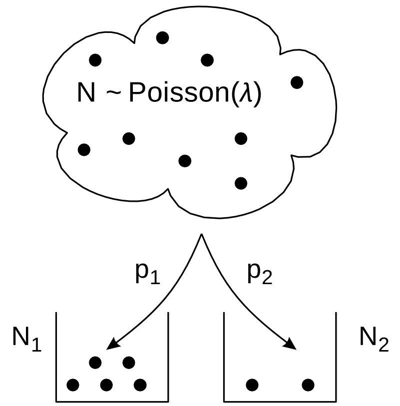
$$G_X(z) = (1-p) + pz, \qquad G_N(z) = e^{(z-1)\lambda}$$

$$G_K(z) = G_N(G_X(z)) = e^{(G_X(z)-1)\lambda} = e^{[(1-p)+pz-1]\lambda} = e^{(z-1)p\lambda}$$

Properties of Poisson distribution (continued)

3. If the elements of a set with size $N \sim \text{Poisson}(\lambda)$ 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

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



Proof: By the law of total probability,

$$\mathrm{P}\{N_1 = n_1, N_2 = n_2\} = \sum_{n=0}^{\infty} \underbrace{\mathrm{P}\{N_1 = n_1, N_2 = n_2 | N = n\}}_{\text{multinomial distribution}} \times \underbrace{\mathrm{P}\{N = n\}}_{\text{Poisson distribution}}$$

$$= \frac{n!}{n_1!\,n_2!} p_1^{n_1} p_2^{n_2} \frac{\lambda^n}{n!} e^{-\lambda} \Bigg|_{n=n_1+n_2} = \frac{p_1^{n_1} p_2^{n_2}}{n_1!n_2!} \lambda^{n_1+n_2} e^{-\lambda \overbrace{(p_1+p_2)}^{1}}$$

$$= \frac{(p_1\lambda)^{n_1}}{n_1!} e^{-p_1\lambda} \times \frac{(p_2\lambda)^{n_2}}{n_2!} e^{-p_2\lambda} = \mathrm{P}\{N_1 = n_1\} \cdot \mathrm{P}\{N_2 = n_2\}$$

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

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

7.5 Generating function of the complementary distribution

Let X be a random variable that assumes integer k with probability p_k and let q_k be the distribution for its tails:

$q_k =$ P{X >k} =$p_{k+1} + p_{k+2} + \cdots$

We denote the PGF of $\{p_k\}$ by P(z) and the generating function of $\{q_k\}$ by Q(z). Then it is not difficult to find the following simple relation:

$$Q(z) = \frac{1 - P(z)}{1 - z}$$

Distributions and Their Relationships