or equivalently
(if x, Y, Z have a joint pdf (or pmf)

$$f_{x,Y,Z}(x,y,z)$$
 so that
 $P((x,Y,Z) \in D) = \iiint f_{x,Y,Z}(x,y,z) dx dy dz$
 $P((x,Y,Z) \in D) = \iiint f_{x,Y,Z}(x,y,z) dx dy dz$
for all regions D,
if $f_{x,Y,Z}(x,y,z) = f_x(x) f_y(y) f_y(z)$
 $T_{x,Y,Z}(x,y,z) = f_x(x) f_y(y) f_y(z)$
for all x, y, z .
Identically distributed rv's have the
same cdf.

Various Facts

$$\begin{split} & \mathsf{E}(\mathsf{X}+\mathsf{Y}+\mathsf{Z}) = \mathsf{E}\mathsf{X} + \mathsf{E}\mathsf{Y} + \mathsf{E}\mathsf{Z} \\ & \text{If } \mathsf{X},\mathsf{Y},\mathsf{Z} \text{ are <u>mutually indep.</u>, then \\ & \mathsf{Var}(\mathsf{X}+\mathsf{Y}+\mathsf{Z}) = \mathsf{Var}(\mathsf{X}) + \mathsf{Var}(\mathsf{Y}) + \mathsf{Var}(\mathsf{Z}) \\ & \mathsf{M}_{\mathsf{X}+\mathsf{Y}+\mathsf{Z}}(t) = \mathsf{M}_{\mathsf{X}}(t) + \mathsf{Var}(\mathsf{Y}) + \mathsf{Var}(\mathsf{Z}) \\ \end{split}$$

If $X_{1}, X_{2}, \dots, X_{n}$ are <u>i.i.d.</u>, then $E\left(\sum_{i=1}^{n} X_{i}\right) = n(EX_{1})$ $Var\left(\sum_{i=1}^{n} X_{i}\right) = n(Var X_{1})$ $M_{n}\left(t\right) = \left[M_{X_{1}}(t)\right]^{n}.$



(2) Var
$$X = EX^{2} - (EX)^{2}$$

Since $X^{2} = X$ and $X^{K} = X$ for all K
since X takes on only the
values 0 and 1 for which
 $1^{K} = 1$ and $0^{K} = 0$.
We have $EX^{2} = p$.
Thus $Var X = p - p^{2} = p(1-p)$.
(3) $M_{\chi}(t) = Ee^{tX} = \sum_{x \in \mathcal{X}} e^{tX} f(x)$
 $= e^{t0} \cdot f(0) + e^{t1} \cdot f(1)$

Think of $X_{1,3}X_{2,3}X_{3,...}$ as the results of repeated trials or tosses of a coin:

 $X_{i} = I_{\{\text{success on trial } i\}}$ or $X_{i} = I_{\{\text{heads on } i^{\text{th}} \text{ toss}\}}.$

 $= 1 - p + pe^{c}$

The Binomial distribution

Let X_1, X_2, \dots, X_n be iid Bernoulli(p). Define $S_n = \sum_{i=1}^n X_i$.

Then $S_n \sim \text{Binomial}(n,p)$

with pmf $P(S_n = K) = {\binom{n}{K}} p^K (1-p)^{n-K}, K=0, 1, ..., n_j$ and

$$\begin{split} & E S_{n} = n E X_{i} = n p , \\ & Var S_{n} = n Var X_{i} = n p (1-p) , \\ & M_{s_{n}}(t) = \left[M_{x_{i}}(t) \right]^{n} = (1-p+pe^{t})^{n} . \end{split}$$

(These results can also be obtained directly from the pmf.)

The Geometric distribution Let X1, X2, X3,... be iid Bernoulli(p). Define T₁ = the # of the trial at which the first success (=1) occurs $= \inf \{ n : X_n = 1 \}.$ Csmallest Then $T_1 \sim \text{Geometric}(p)$, $P(T = K) = (1 - p)^{K-1} p$ for K=1,2,3,... $ET_{i} = \frac{1}{p}$, $Var T_{i} = \frac{1-p}{p^{2}}$, $M_{T_{1}}(t) = \frac{pe^{t}}{1 - (1 - p)e^{t}}, t < -\log(1 - p).$ There is no guick way to "see" the mean, variance and mgf. We will obtain the mgf from the pmf, and then use the mgf to get the mean and variance.

The mgt Let X~ Geometric (p). $M_{x}(t) = Ee^{tX} = \sum_{x} e^{tx} f_{x}(x)$ $x \in \mathcal{Y}$ $= \sum_{x=1}^{\infty} e^{tx} (1-p)^{x-1} p$ $(e^{t})^{x} = e^{t} \cdot (e^{t})^{x-1}$ $= pe^{t} \sum_{x=1}^{\infty} [(1-p)e^{t}]^{x-1}$ A geometric series which converges if (1-p)et<1, or equivalently t<-log(1-p) $\frac{pe^{t}}{1-(1-p)e^{t}} \quad \text{for } t < -\log(1-p)$ (undefined otherwise)















The Discrete Memoryless Property

(1) If X~ Geometric (p), then X satisfies the Discrete Memoryless Property (DMP): P(X>y+z|X>y) = P(X>z) for all integers y, z>0, or equivalently,

$$P(X>y+z) = P(X>y)P(X>z) \quad (\mathbf{f})$$

for all integers $y, z > 0$.

Conversely,
(2) If the rVX has range
$$\mathcal{X} = \{1, 2, 3, ...\}$$

and X satisfies the DMP,
then X~ Geometric (p) for some value of p.

Proof of (1):

$$P(X > y + z | X > y) = \frac{P(\{X > y + z\} \cap \{X > y\})}{P(X > y)}$$

$$= \frac{P(X > y + z)}{P(X > y)} (*)$$

Here we have used the fact that

$$\{x > y + z\} \subset \{x > y\}$$
so that
$$\{x > y + z\} \cap \{x > y\} = \{x > y + z\}$$

$$\begin{bmatrix} \text{In general: If event A implies event B,} \\ \text{then ACB so that A \cap B = A.} \end{bmatrix}$$
Now note that $P(x > x) = (1 - p)^{x}$, $x = 1, 2, 3, ...$
Plugging this into (*) yields
$$\frac{P(x > y + z)}{P(x > y)} = \frac{(1 - p)^{y + z}}{(1 - p)^{y}} = (1 - p)^{z}$$

$$= P(x > z).$$
Proof of (2): Versian (†) of DMP with
$$z = 1 \text{ gives}$$

$$P(x > y + 1) = P(x > y) P(x > 1) \text{ for } y = 1, 2, 3, ...$$
Setting $y = 1 \text{ gives}$

$$P(x > 2) = P(x > 1)^{z}.$$
Then $y = 2 \text{ gives}$

$$P(x > 3) = P(x > 2) P(x > 1) = P(x > 1)^{3}$$

and by induction we see $P(X > K) = P(X > 1)^{K} \text{ for } K = 1, 2, 3, \dots$ Taking p = 1 - P(X > 1) this becomes $P(X > K) = (1 - p)^{K} \text{ or}$ $P(X \le K) = 1 - (1 - p)^{K} \text{ for } K = 1, 2, 3, \dots$ which is the cdf of a Geometric (p) distn. For a discrete random variable X and event A, define

$$E(X \mid A) = \sum_{x \in \mathcal{X}} x P(X = x \mid A)$$
$$E(g(X) \mid A) = \sum_{x \in \mathcal{X}} g(x) P(X = x \mid A)$$

whenever the sums converge absolutely.

Analog of "Law of Total Probability" for expectations: If the events B_1, B_2, \ldots, B_k form a partition, then

$$Eg(X) = \sum_{i=1}^{k} E(g(X) \mid B_i) P(B_i).$$

Note: It is also true when $k = \infty$. Proof:

$$\sum_{i=1}^{k} E(g(X) | B_i) P(B_i)$$

$$= \sum_{i=1}^{k} \left(\sum_{x \in \mathcal{X}} g(x) P(X = x | B_i) \right) P(B_i)$$

$$= \sum_{x \in \mathcal{X}} \sum_{i=1}^{k} g(x) P(X = x | B_i) P(B_i)$$

$$= \sum_{x \in \mathcal{X}} g(x) \sum_{i=1}^{k} P(\{X = x\} \cap B_i)$$

$$= \sum_{x \in \mathcal{X}} g(x) P(X = x) = Eg(X).$$

Conditional expectations have the same properties as ordinary expectations. For example:

E(aX + bY + c | A) = aE(X | A) + bE(Y | A) + c

Application: Mean and Variance of Geometric Distribution:

Suppose $X \sim \text{Geometric}(p)$.

The Discrete Memoryless Property can be written: For all y > 0,

 $P(X - y > z \mid X > y) = P(X > z)$ for z = 1, 2, 3, ...

Or in other words:

 $\mathcal{L}(X - y \mid X > y) = \mathcal{L}(X) \text{ which implies}$ $E(g(X - y) \mid X > y) = E(g(X)) \text{ for any } g$ $(``\mathcal{L}'' \text{ means ``distribution'' or ``law'')}$

Mean of *X*:

$$EX = E(X | X = 1)P(X = 1) + E(X | X > 1)P(X > 1)$$

= $1 \cdot p + E(1 + (X - 1) | X > 1) \cdot (1 - p)$
= $p + E(1 + X) \cdot (1 - p)$
= $p + (1 + E(X))(1 - p)$ so that
 $pEX = 1$ leading to $EX = 1/p$.

Variance of *X*:

Similarly,

$$EX^{2} = E(X^{2} | X = 1)P(X = 1) + E(X^{2} | X > 1)P(X > 1)$$

$$= 1 \cdot p + E\{(1 + (X - 1))^{2} | X > 1\} \cdot (1 - p)$$

$$= 1 \cdot p + E\{(1 + X)^{2}\} \cdot (1 - p)$$

$$= p + (1 + 2EX + EX^{2})(1 - p)$$

$$= p + \left(1 + \frac{2}{p} + EX^{2}\right)(1 - p) \text{ so that}$$

$$pEX^{2} = \frac{2}{p} - 1 \text{ and } EX^{2} = \frac{2 - p}{p^{2}}$$

Thus

Var(X) =
$$EX^2 - (EX)^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$$
.

Additional Exercise: Compute EX^3 by this approach.

The Negative Binomial Distribution

Let X1, X2, X3,... be iid Bernoulli(p). If $T_r =$ the time of the rth success = the number of trials needed to obtain r successes $= \inf f n : S_n = r$ where $S_n = \sum_{i=1}^n X_i$, then Tr~ Negative Binomial (r,p) with pmf $P(T_r = x) = \begin{pmatrix} x - i \\ r - i \end{pmatrix} p^r (i - p)^{x - r}$ for $\chi = r_{1}r_{1}r_{1}r_{2}...$ $ET_r = \frac{r}{p}$ $Var T_r = r \left(\frac{1-p}{p^2} \right)$ mgf $M_{T_r}(t) = \left(\frac{pet}{1-(1-p)et}\right)^r$ for $t < -\log(1-p)$

Argument: $T_{r} = T_{1} + (T_{2} - T_{1}) + \cdots + (T_{r} - T_{r-1})$ since the coin has no memory, it is intuitively obvious that $T_1, T_2 - T_1, T_3 - T_2, \dots, T_r - T_{r-1}$ are iid Geometric (p). Therefore $ET_r = r(ET_i) = \frac{r}{p}$ $\operatorname{Var} T_r = r(\operatorname{Var} T_i) = r(\frac{1-p}{p^2})$ $M_{T_r}(t) = (M_{T_1}(t))^r = (pet_{(1-p)et})^r$ for t < - log (1-p) Derivation of pmf: ${T_r = k} = {$ the kth trial is a success and exactly r-1 of the

previous trials are successes}

$$\{T_r = k\} = \{X_k = l\} \cap \{S_{k-1} = r-l\}$$
where $S_n = \#$ of successes in $n \text{ trials},$
and $X_n = \text{result of } n^{\text{th}} \text{ trial}.$
Clearly $\{X_k = l\}$ and $\{S_{k-1} = r-l\}$

are independent and $S_{K-1} \sim \text{Binomial}_{(k-1,p)}$. Thus (for $K \ge r$)

Thus (for
$$R \ge r$$
)

$$P(T_r=K) = P(X_K=I) P(S_{K-I}=r-I)$$

$$= p \binom{k-I}{r-I} p^{r-I} (I-p)^{k-r}$$

$$= \binom{k-I}{r-I} p^r (I-p)^{k-r}$$

Note: My definition of the negative binomial is the book's "atternate definition".

Given a sequence of independent coin tasses,
each with probability
$$p$$
 of heads, define
 $S_n = \#$ of heads in n tosses,
 $T_r = \#$ of tosses needed to get r heads.
We know that
 $S_n \sim \text{Binomial}(n,p)$,
 $T_r \sim \text{Neg.Bin}(r,p)$.
Finding the cdf of T_n
() Direct derivation (from defn. of cdf).
Recall $P(T_r = \mathring{J}) = (\mathring{J}_{r-1}^{-1})p^r(i-p)\mathring{J}_{r}^{-r}$
for $\mathring{J} = r, r+1, r+2, \cdots$
Thus $P(T_r \leq K) = \sum_{j=r}^{K} (\mathring{J}_{r-1}^{-1})p^r(i-p)\mathring{J}_{r}^{-r}$ (*)
 \mathring{J}_{r}^{-r} integers $K \geq r$, and
 $P(T_r \leq K) = 0$ for $K < r$.

2) Indirect derivation
Use the property

$$\{T_r > n\} = \{S_n < r\}.$$

This holds for all positive integers r and n.
Thus, for any positive integer K,
 $P(T_r \le K) = I - P(T_r > K) = I - P(S_k < r).$
Now
 $P(S_k < r) = 1$ if $K < r$, and
 $= \sum_{j=0}^{r-1} {K \choose j} p^j (1-p)^{K-j}$ if $K \ge r.$

Thus $P(T_{r} \leq \kappa) = I - \sum_{j=0}^{r-1} {\binom{k}{j}} p^{j} (I-p)^{k-j} \text{ for } k \geq r$ $= 0 \quad \text{for } k < r.$

This result and the earlier formula (*) give different formulas for the same guantity. For small r, this second formula is usually easier to use. Closure" Properties Suppose X1, X2 are independent. 1) If X,~Binomial(N1,p) and $X_2 \sim \text{Binomial}(n_2, p)$ then $X_1 + X_2 \sim \text{Binomial}(n_1 + n_2, p)$. 2 If XIN Neq. Bin. (ri,p) and $X_2 \sim Neq. Bin. (r_2, p),$ then $X_1 + X_2 \sim Neq. Bin. (r_1 + r_2, p)$. [Recall: Geometric (p) same as NB(1,p).] similar properties hold for sums of three or more rus. These properties are intuitive. Give a "coin tossing" story for each. Proofs: Use mgf's. Recall that mgf's are unique, and that $M_{X_1+X_2}(t) = M_{X_1}(t)M_{X_2}(t)$ if X, and X2 are independent.

 $(I) \quad M_{X_1+X_2}(t) = M_{X_1}(t) M_{X_2}(t)$ $= (1 - p + pe^{t})^{n_{1}} (1 - p + pe^{t})^{n_{2}}$ $= (1 - p + pe^{t})^{n_1 + n_2}$ = mgf of Binomial (n_1+n_2, p) . Thus (by uniqueness) $X_1 + X_2 \sim Binomial(n_1 + n_2, p)$. (2) $M_{X_1+X_2}(t) = M_{X_1}(t) M_{X_2}(t)$ $= \left(\frac{pe^{t}}{1-(1-p)e^{t}}\right)^{r_{i}} \left(\frac{pe^{t}}{1-(1-p)e^{t}}\right)^{r_{2}}$ $= \left(\frac{pe^{t}}{1-(1-p)e^{t}}\right)^{r_1+r_2}$ = mgf of Neg.Bin (r_1+r_2, p) . Thus (by uniqueness of mgf's) $X_1 + X_2 \sim Neq. Bin(r_1 + r_2, p).$

Hypergeometric Distribution (arises when sampling from a finite population without replacement) Suppose you have an urn with: R red balls, G green balls. Draw k balls at random (without replacement). Define $S_{K} = \# of red balls in sample.$ The pmf of Sk is $P(S_{k} = x) = \begin{pmatrix} \# \text{ of samples of size } k \\ containing exactly x \\ red balls \end{pmatrix}$ (# of samples of size K)

 $= \frac{\binom{R}{\binom{G}{K-\chi}}\binom{G}{K-\chi}}{\binom{R+G}{K}} \quad for \quad 0 \le \chi \le R$

(pmf = 0 otherwise.)

SK has a hypergeometric distribution. In the textbooks notation: $S_{k} \sim Hypergeometric (N = R + G, M = R, K = k)$ The mgf of SK $S_{K} \in \{0, 1, 2, ..., k\}$ Since Sk is bounded, the mgf' is finite for all t. But the mgf has no simple closed form. (There is no way to simplify the summation.) 50 mgf is not useful. Mean and Variance of SK The book computes these directly from the pmf. We use another route.

Define
$$X_i = \begin{cases} 1 & \text{if } i^{\text{th ball is red}}, \\ 0 & \text{otherwise} \end{cases}$$

= $I_{\{i^{\text{th ball is red}\}}.$
Then $S_k = \sum_{i=1}^{k} X_i.$
Since $EX_i = P(i^{\text{th ball is red}})$
= $P(1^{\text{st ball is red}}) = \frac{R}{R+G},$
(see discussion below)

we get $ES_{k} = E\sum_{i=1}^{k} X_{i} = \sum_{i=1}^{k} EX_{i} = K \frac{R}{R+G}$. <u>Discussion</u>: We are now thinking of the sample of k balls (chosen from R+G) as being ordered. There are $(R+G)(R+G-1)\cdots(R+G-k+1)$ $= \binom{R+G}{K} K!$ equally likely ordered samples. The number of ordered samples where the *i*th ball is red is

$$R \cdot {\binom{R+G-1}{K-1}} \cdot {\binom{K-1}{K}}$$

$$p_{ick one} \qquad \uparrow \qquad \uparrow$$

$$p_{icd ball} \qquad p_{ick the} \qquad p_{lace them in}$$

$$for the \qquad other K-1 \qquad the remaining$$

$$i^{th} position \qquad balls \qquad K-1 positions$$

$$Thus \\P(i^{th} ball is red) = \qquad \frac{R \cdot {\binom{R+G-1}{K-1}} \cdot {\binom{K-1}{K}}}{{\binom{R+G}{K}} K!}$$

$$= \frac{R \frac{(R+G-1)!}{(K-1)!(R+G-K)!} (K-1)!}{\frac{(R+G)!}{K!(R+G-K)!}} = \frac{R(R+G-1)!}{(R+G)!}$$

$$= \frac{R}{R+G} \cdot$$
We will shortly need
For $i \neq j$,
$$P(i^{th} and j^{th} balls are red)$$

$$= P(1^{st} and 2^{nd} balls are red)$$

$$= \frac{R}{R+G} \cdot \frac{R-1}{R+G-1}$$

A counting proof: The number of ordered samples where the ith and ith balls are red is so that P(ith and ith balls are red) $R(R-1)\binom{R+G-2}{k-2}(K-2)!$ $\binom{\mathsf{R}+\mathsf{G}}{\mathsf{k}}$ k! $= \frac{R(R-1)(R+G-2)!}{(R+G)!} = \frac{R \cdot (R-1)}{(R+G) \cdot (R+G-1)} \cdot$



Note that

$$X_{i}X_{j} = I_{\{i^{th} ball red\}}I_{\{j^{th} ball red\}}$$

$$= I_{\{i^{th} and j^{th} ball are red\}}$$
(recall $I_{A}I_{B} = I_{A \cap B}$.)
so that
 $E X_{i}X_{j} = P(i^{th} and j^{th} balls are red)$
(recall $E I_{C} = P(C)$)

$$= \int \frac{R}{R+G} \cdot \frac{R-1}{R+G-1} \quad \text{for } i \neq j$$
($\frac{R}{R+G} \quad for i = j$.)
Thus $ES_{k}^{2} = \sum_{i=1}^{k} \sum_{j=1}^{k} E X_{i}X_{j}$

 $= K \frac{R}{R+G} + K(K-1) \frac{R}{R+G} \cdot \frac{R-1}{R+G-1} \cdot$

$$Var(S_{k}) = ES_{k}^{2} - (ES_{k})^{2}$$

$$= K \frac{R}{R+G} + K(k-1) \frac{R}{R+G} \cdot \frac{R-1}{R+G-1}$$

$$- (K \frac{R}{R+G})^{2}$$

$$= K \frac{R}{R+G} \left[\frac{(R+G)(R+G-1) + (K-1)(R-1)(R+G)}{-K R(R+G-1)} \right]$$

$$= K \frac{R}{R+G} \cdot \frac{G}{R+G} \cdot \frac{R+G-K}{R+G-1}$$

$$Variance of$$
Binomial $(n=K_{3}p=\frac{R}{R+G})$
finite population correction factor

The Poisson Distribution
If
$$X \sim Poisson(\lambda)$$
, then
 $pmf P(X=k) = \frac{\lambda^{K}e^{-\lambda}}{K!}$ for $k=0,1,2,...$
 $EX = Var X = \lambda$
 $mgf M_{X}(t) = e^{\lambda(e^{t}-1)}$ (exercise)
Closure Property
If $X_{1} \sim Poisson(\lambda_{1}), X_{2} \sim Poisson(\lambda_{2}), and$
 X_{1} and X_{2} are independent,
then $X_{1} + X_{2} \sim Poisson(\lambda_{1} + \lambda_{2})$.
Proof:
 $M_{X_{1}} + X_{2}(t) = M_{X_{1}}(t) M_{X_{2}}(t)$ by independence
 $= e^{\lambda_{1}(e^{t}-1)} \cdot e^{\lambda_{2}(e^{t}-1)}$
 $= e^{(\lambda_{1} + \lambda_{2})(e^{t}-1)}$
 $= mgf of Poisson(\lambda_{1} + \lambda_{2})$

The Poisson distr. arises because of Poisson approximation to the Binomial distn. If X~Binomial(n,p) where n is large and p is small, then $P(X=k) \approx \frac{\lambda^{k}e^{-\lambda}}{\kappa_{l}}$ with $\lambda = np$ (that is, $X \sim \text{approx Poisson}(\lambda = np)$). Formal statement as a limit theorem Fix a value 2>0. Suppose $X_n \sim \text{Binomial}(n, p = \frac{\lambda}{n})$ for n = 1, 2, 3, ...and $Y \sim Poisson(\lambda)$. Then $\lim_{n \to \infty} P(X_n = K) = \frac{\lambda^{\kappa} e^{-\lambda}}{\kappa_1} = P(Y = K),$ that is, $X_n \xrightarrow{d} Y$ (convergence in distr.) Poisson approximation to hypergeometric distr. If X~hypergeometric(N, M, K) and T [# of red balls total # # of red balls] in sample of K of balls in urn N is large, M is small, and K is small, then $X \sim approx Poisson(\lambda = k \cdot \frac{M}{N})$.

More general Poisson approximation

Suppose $A_1, A_2, ..., A_n$ are <u>independent</u> events and $Z_1, Z_2, ..., Z_n$ are the corresponding indicator random variables $Z_i = I_A = \begin{cases} 1 & \text{if } A_i \text{ occurs} \\ 0 & \text{otherwise.} \end{cases}$ Let $p_i = P(A_i)$. If n is large and all the p_i are small, then $S = \sum_{i=1}^{n} Z_i$

has approximately a Poisson $(\lambda = \sum_{i=1}^{n} p_i)$ distn.

Examples

- ① Geiger counter with lump of radioactive material consisting of different isotopes.
- (2) Traffic accidents (on a certain highway during a given period) with drivers of differing ability.
- ③ Number of cases of a rare disease (in a given city) when people vary in their susceptibility.

etc.

Example: Suppose
2000 skiers at a resort,
each has probability .002 of an
accident on any given day,
skiers are independent.
Then

$$X = \# \text{ of accidents today}$$

 $\sim \text{Binomial}(n = 2000, p = .002)$
 $(large)$ (small)
 $\sim \text{approx Poisson}(\lambda = 2000(.002) = 4).$
Thus $P(X = 2) \approx \frac{\lambda^2 e^{-\lambda}}{2!} = \frac{4^2 e^{-4}}{2}.$
Continuation: Similar situation but now
 $P(\text{skier } i \text{ has accident}) = P_i = 2 \times 10^{-6} i$
for $i = 1, 2, \dots, 2000.$
Maximum $P_i = P_{2000} = 2 \times 10^{-6} \times 2000$
 $= 4 \times 10^{-3}$ which is "small".
Thus $X \sim \text{approx Poisson}(\lambda = \frac{2000}{2} p_i).$
 $\lambda = 2 \times 10^{-6} \sum_{i=1}^{2000} i = 2 \times 10^{-6} \cdot \frac{2000 \cdot 2001}{2} \approx 4$
so $P(X = 2) \approx \text{ same as before.}$

The Poisson Process

A process of arrivals (clicks, events, etc.) is called a Poisson process with constant rate λ if

 The number of arrivals in any given fixed period of time of duration t has

 Poisson (λt) distribution.

 Disjoint intervals of time are independent.

Consequence of (1): The expected number of arrivals in a period of length t is λt . The arrivals occur at rate λ (on average).

Examples:

 Clicks on a Geiger counter (under uniform conditions)
 Accidents on a highway (under uniform conditions)

etc.

Example (Poisson Process) Clicks on a Geiger counter occur according to a Poisson process with average rate 1.5 clicks per second. what is the probability of exactly 2 clicks' during (0,3) and exactly 3 clicks during (2,6)? These time periods overlap. So we do the following. Let $X_1 = #$ of clicks during (0,2), $X_2 = #$ of clicks during (2,3), disjoint intervals $X_3 = #$ of clicks during (3, 6).] of time Then X1, X2, X3 are independent with \times , ~ Poisson ($\lambda_1 = 1.5 \times 2 = 3$) $X_2 \sim Poisson (\lambda_2 = 1.5 \times 1 = 1.5)$ $X_3 \sim Poisson (\lambda_3 = 1.5 \times 3 = 4.5).$ We want $P(X_1 + X_2 = 2, X_2 + X_3 = 3)$ $= P(X_1 = 2, X_2 = 0, X_3 = 3)$ + $P(X_1 = 1, X_2 = 1, X_3 = 2)$ + $P(X_1 = 0, X_2 = 2, X_3 = 1)$





 $= \left(\frac{9 \cdot 4.5^{3}}{12} + \frac{3 \cdot 1.5 \cdot 4.5^{2}}{2} + \frac{1.5^{2} \cdot 4.5}{2}\right)e^{-9}$