1 CH1. Distribution, P-Proc, Nonhomog P-Proc

Dist.: $Poi(\mu)$ parameter μ given by $p_k = e^{-\mu}\mu^k/k!$, $k \in \mathbb{N}_0$. $E = \mu$, $EX^2 = \mu^2 + \mu$. Unusual characteristic: mean = variance.

Noteworthy theorems. $X + Y$ Pois parameter the sums of both parameters. Second: let $N \sim Poi(\mu)$; and conditional on N, let $M \sim Bin(N, p)$. Then unconditional dist of M is $Poi(\mu p)$.

P-Proc. Notions entail independence and Pois dist. Intensity, or rate, $\lambda > 0$ is int-valued stoc process $\{X(t); t \geq 0\}$:

- For $t_0 < t_1 < t_2$, increments $X(t_1) - X(t_0)$, $X(t_2) - X(t_1)$ indep r.v.,
- $s \geq 0, t > 0$, r.v. $X(s+t) - X(s)$ Pois dist p_k has parameter $\mu = \lambda \cdot t$,
- $X(0) = 0$. Note that $P(X(k) = 0) = e^{-\lambda k}$.

Nonhomog Proc. Above, rate λ is proportionality constant in the probability of event occuring at small intervals (that is, $P(X(t + h) - X(t) = 1) \sim \lambda h + o(h)$ as $h \to 0$. Note here that $o(h) \subset O(h)$.

Considering rates $\lambda = \lambda(t)$ that vary in time (we do not consider this often), if $X(t)$ nonhomog with such a rate, then $X(t) - X(s)$ Pois dist parameter (integral s to t $\lambda(u)du$).

2 What is this section I don't understand

Overview? This law asserts that [when certain event may occur in any of large possibilities, but probability of event occur in any given possibility is small, then total num of events happen should approximately follow Pois]

Theorem 1.2.1 Let ϵ_i indep Bernoulli r.v. with $P(\epsilon_i = 1) = p_i$, and S_n = sum of n of those. Then exact probability (distribution) for S_n and (prob dist) of P-Proc with μ = sum of p_i satisfies $\left|P(S_n = k) - \mu^k e^{-\mu}/k!\right| \leq \sum_{i=1}^n p_i^2$. This means the Pois dist provides good approx to exact probability when p_is uniformly small.

Proof roughly follows: separate $X(\mu)$ to sum of indep $X(p_i)$ and compare pointwise to obtain loose upper bound. Counting $P(\epsilon(p) = X(p)) = 1 - p + pe^{-p}$ turns out to be $\geq 1 - p^2$.

This turns out to be related to P-Procs? Let $N((a, b])$ equal to num of events in that interval (imagine events on $\mathbb{R}^+$ but "restricted"to that interval). Postulates are 1. # disjoint intervals are indep, 2. For any t, h the dist of $N((t, t+h])$ depends only on h, 3. Exist positive λ that $P(N((t, t + h]) \geq 1) = \lambda h + o(h)$ as $h \downarrow 0$. [events are rare], 4. $P(N((t, t + h]) \geq 2) = o(h)$ as $h \downarrow 0$ [excludes the possibility of simultaneous occurence of 2 or more events].

Rough proof: divide into n subintervals of equal length, and set $\epsilon_i = 1$ if exist at least one event in interval i.

Definition. This motivates $N((s, t])$ r.v. counting num of events in that interval; it is P-Point-Proc intensity λ if 1. disjoint intervals independent, 2. for any $s < t$, $N((s, t])$ has Pois dist $P(N = k) = [\lambda(t - s)]^k e^{-\lambda(t-s)}/k!$ for $k \in \mathbb{N}_0$.

3 Waiting Times Distribution

Some semantics. P-Point-Proc $N((s, t])$ counts num of events in interval, P-Proc $X(t)$ num of events occuring until t (or, $= N((0, t])$). Define waiting time W_n be time of occurence of nth event, and sojourn times is $S_n = W_{n+1} - W_n$.

Theorems 1.3.1-3.

- W_n has gamma dist with pdf $\frac{\lambda^n t^{m-1}}{(n-1)!}e^{-\lambda t}$, $n \geq 1$. In particular, W_1 exponentially distributed ($= \lambda e^{-\lambda s}$).
- Sojourn times indep r.v., each exponential (as above).
- $P(X(u) = k|X(t) = n) = nCk(u/t)^k(1 - (u/t))^{n-k}$.

4 Major result of Uniform Dist and examples

Introduction, and some experiment. The result asserts that, conditioned on fixed total num of events in an interval, locations of those events are uniformly distributed ("in a certain way"???)

Experiment: begin with line length t units, throw fixed n darts (uniformly thrown) and indep of location of other darts. U_i be position of ith dart, and $f_U(u) = 1/t$ for that segment, 0 elsewhere.

Now let $W_1 \leq \cdots \leq W_n$ denote same positions, with order also in consideration. We fix NOT the endpoint of an interval (this is a more special case of this theorem), BUT that exactly n events happens before/equal a certain point.

Joint pdf is $f_{W_1,\ldots,W_n|X(t)=n}(w_1, \ldots, w_n) = n!t^{-n}$.

I will rewrite the proof this time. We calculate the multiplication of "complement of small intervals all 0" and "each small interval, one event occurs".

First is $P(N((0, w_1]) = 0, \ldots, N((w_n + \Delta w_n, t]) = 0)$, equals $e^{-\lambda t}e^{\lambda(\Delta\text{sum deltas})} = e^{-\lambda t}(1 + O(\text{max delta}))$.

Second is $P(\text{all } N = 1)$ equals $(\prod(\lambda\Delta w_i))(1 + O(\text{max delta})$

Thus value of f is

$$
\begin{aligned}
&f_{W_i|X(t)=n}(w_i)\prod \Delta w_i \\
=&P(w_i < W_i \leq w_i + \Delta w_i|X(t) = n) + o(\prod \Delta w_i) \\
=&\frac{e^{-\lambda t}\prod(\lambda\Delta w_i}{e^{-\lambda t}(\lambda t)^n/n!}(1 + O(\max \Delta w_i)) + o(\prod \Delta w_i).
\end{aligned}
$$

where last line follows by Bayes expansion. Dividing by prod delta and letting delta to 0 gives result.

Immediate Example 1. Customers arrive according to P-Proc rate λ. Each pays 1 on arrival, and want to evaluate expected value "discounted back to time 0". Quantity is $M = E\left(\sum_{i=1}^{X(t)} e^{\beta W_i}\right)$, where β is the discount rate.

Evaluate by conditioning on $X(t) = n$. This M is now equal to $\sum_{i=1}^{\infty} E(\text{above}|X(t) = n)P(X(t) = n)$. Due to symmetry (then applying) earlier theorem, we have the first part(s) to be= $E\left(\sum_{k=1}^n e^{\beta U_k}\right) = n/(\beta t)(1 - e^{\beta t})$.

Substitution gives M = earlier$/nE(X(t))$ (also due to $\sum nP(X(t) = n) = EX(t)$).

Immediate Example 2. Viewing fixed mass of radioactive material, alpha particles appear P-Proc λ. Each particle exists for random duration then annihilated.

Successive lifetimes Y_i of distinct particles are indep, with distinct fn $G(y) : P(Y_i \leq y)$. Let $M(t)$ = # existing particles at time t.

- Clearly $M(t) \leq X(t)$. We condition on $X(t) = n$.
- Particle k exists at t if $W_k + Y_k \geq t$.
- $P(M(t) = m|X(t) = n)$ equals to sum of indicator variables equal to m;
- Symmetry among particles gives $P(\sum \text{indic var} = \text{m}|X(t) = n)$ equals to $P(\sum I(U_k + Y_k \geq t) = m)$.

Note that particle k exists at t iff $W_k + Y_k \geq t$.

After calculating $p = P(U_k + Y_k \geq t)$ which is equal to $1/t\int_0^t(1 - G(z))dz$, and $P(M = m|X = n) = nCm \cdot p^m(1 - p)^{n-m}$, we get $M(t)$ is Pois distributed with mean λpt.

5 Shot Noise and Sum Quota Sampling

Shot Noise. Model for current fluctuations due to chance arrivals of electrons to anode.

- Arrive according to P-Proc λ, arriving electron provides current whose intensity after x time is impulse function $h(x)$.
- Intensity of current at t is shot noise $I(t) = \sum_{k=1}^{X(t)} h(t - W_k)$.
- We show that $I(t)$ has same dist as certain random sum: define U_i iid uniform, define $\epsilon_k = h(U_k), k \geq 1$. Claim $I(t) = S(t) := \sum_{i \leq X(t)} \epsilon_i$.

NOTE: derivation of LOTV to something "applicable when $S = \sum_{i \leq N} X_i$, N is r.v."

$$Var(S) = E_N(Var_{X_i}(S|N)) + Var_{X_i}(E_N(S|N))$$

equal to $E(N)Var(X_i) + E(X_i)^2 Var(N)$ (corresponding to each of the two terms above). Simplified, it becomes $E(N)E(X_i^2) + E(X_i)^2(Var(N) - E(N))$.

We can then get $EI(t) = \lambda\int_0^t h(u)du$ and $Var(I(t)) = \lambda tE(h^2(U_1))$.

Proof of theorem: $P(I \leq t)$, by the same maneuver, equals to $\sum P(\sum h(t - W_k) \leq x)P(X(t) = n)$. This is equal to $\sum P(\sum \epsilon_k \leq x)P(X(t) = n)$, giving our desired result.

SQS. Estimate mean service life of a part, given a preassigned TIME quota t. Define $N(t)$ is max n such that $\sum_{i \leq n} X_i < t$. The "natural" sample mean $\overline{X}$ is $\sum_{i \leq n} X_i/N(t)$. Note that we must assume $N(t) > 0$.

This is difficult, but can be carried out when X exponential distribution with common parameter λ.

Key is use theorem to evaluate that $E(W|N(t) = n)$ is $tn/(n + 1)$, and $E(W/N|N(t) > 0)$ is $1/\lambda(1 - \lambda t/(e^{\lambda t} - 1)$.

Expressing this as a ratio to the mean bias $EX_1 = 1/\lambda$, it equals $\frac{EX1 - E\overline{X}}{EX1} = \lambda t/e^{\lambda t} - 1$. Name the fraction that solely depends on λt a special term called "fraction bias". The values are $(1, 2, 3, 4, 5, 6, 10)$ to $(0.58, 0.31, 0.16, 0.07, 0.03, 0.015, 0.0005)$.

Example: if $N(t) = 3$ and sample mean is $\overline{X} = 7$, a more accurate estimate is $7/0.84$ (sample/(1-fraction bias)).

6 Spatial P-Proc

Define multidimensional P-Proc and describe examples in this section. Definitions.

- Set S to be a set in n-dimensional space, and $\mathcal{A}$ a family of subsets of S.
- Point process in S is stoc process $N(A)$ indexed by $A \in \mathcal{A}$, having nonnegative integers as possible values.
- This "means" to count points "resulting inside A" as $N(A)$. So, if A and B disjoint and in $\mathcal{A}$ and $A \cup B \in \mathcal{A}$, then $N(A \cup B) = N(A) + N(B)$.

The 1D case of $S = \mathbb{R}^+$ and $\mathcal{A} = \{(s, t]\}$ is introduced in Section 1.3 (note: where $N((0, t]) = X(t)$). This 3D version has relevance when considering (stars, astronomy), (wildlife, ecology), (bacteria, medicine), (defects, surface) and (reliability, volume).

Let $|A|$ size (length, area, volume) of A then $\{N(A) : A \in \mathcal{A}\}$ is homog P-Point-Proc of rate $\lambda > 0$ if

- for $A \in \mathcal{A}$, r.v. $N(A)$ is Pois dist with parameter $\lambda|A|$;
- for finite collection of disjoint subsets, r.v. of those are indep.

The multidimensional case is nalogous to Section 1.2, law of rare events are invoked to derive Poisson process as a consequence of certain postulates. The postulates are:

- If $0 < |A| < \infty$ then probability is strictly between 0 and 1. The r.v. takes nonnegative values.
- Dist of $N(A)$ is $\lambda|A| + o(|A|)$ when $|A| \to 0$. This only depends on A's size.
- Disjoint implies independence, and r.v. of union is sum of r.v. (this is a copy of the second item in the previous itemize).
- $\lim_{|A| \to 0}$ of $P(N(A) \geq 1)/P(N(A) = 1)$ is 1.

Note that if $N(A) = 1$, the event occurs uniformly in A, with regards to area. Proof is just Bayes rule, defined on $C := A\backslash B$

Astronomy example. Let $\mathbf{x}, \mathbf{y}$ 3D vectors, and let light intensity exerted at $\mathbf{x}$ by a star in $\mathbf{y}$ is $f(\mathbf{x}, \mathbf{y}, \alpha)$ equals $\alpha/\|\mathbf{x} - \mathbf{y}\|^2$, and α a r.v. depending on "star intensity".

Assume α in different stars iid with mean μ_α and var σ_α^2. Also assume combined/accumulated intensity (let it to be $Z(\mathbf{x}, A)$, A any set, and stars is "random") is additive.

This Z is equal to $\sum_{r \leq N(A)} \alpha_r/\|\mathbf{x} - \mathbf{y_r}\}^2$. The expectation of Z can be broken down to $E(N(A))$ and $E(f)$, further broken down to $E\alpha$ and E of denominator.

This is equal to $1/|A|$ times integral of denominator $d\mathbf{y}$ over the set A (tbh alr forgot how to evaluate this).

7 Compound and Marked P-Proc

Preliminaries. P-Proc rate λ, and each event associated with r.v. (e.g. a value or cost). The values of rv Y_i assumed independent (among themselves and the P-Proc). They share common distribution $G(y) = P(Y \leq y)$.

Compound P-Proc: $Z(t) = \sum_{k \leq X(t)} Y_k$, $t \geq 0$. Marked P-Proc: Sequence of pairs (W_i, Y_i), W_i waiting time/event times.

Compound. $E(Z(t)) = \lambda\mu t$, $Var(Z(t)) = \lambda(Var(Y_1) + \mu^2)t$.

Note: calculate $Z(t)$ can be explicit by conditioning on $X(t)$. Recall $G^{(n)}(y) = P(\sum Y_i \leq y)$ is equal to $\int_{-\infty}^{\infty} G^{(n-1)}(y - z)dG(z)$; it is different from regular convolution; Gs denote cumulative function, dGs dempte "slope of G at z/instantaneous rate at z". Consider the last value at z, it is the pdf, and multiplied by earlier Y_is' matching value.

Marked. ACTUALLY TOO LAZY, not much different from the rest; maybe just revisit for practice later.

8 CH2. Pure-Birth Process (finding solution to 1 set of equations)

The equations/postulates.

- $P(X(t+h) - X(t) = 1 | X(t) = k) = \lambda_k h + o_{1,k}(h)$ as $h \to 0+$.
- Prob of equal population is $1 - \lambda_k h + o_{2,k}(h)$.
- Prob of negative change is 0, and $X(0) = 0$.

The functions $o_{i,k}(h)$ not depend on t by "stationarity" of "Markov process". Also note that $X(t) \neq$ population size, BUT equal to # births in $(0, t]$.

Analyzing the possibilities at time t "just prior" to time $t+h$ (invoking LOTP and M-Property; there is half-a-page of conditioning "from 0 to $n-2$, with the significant cases only at $n-1, n$") to prove the differential properties/equations of

- $P_0'(t) = -\lambda_0 P_0(t)$;
- $P_n'(t) = -\lambda_n P_n(t) + \lambda_{n-1} P_{n-1}(t)$, $n \geq 1$.

We have $P_0 = e^{-\lambda_0 t}$ and $P_n(t)$ is prob that t between $\sum_{i \leq n-1} S_i$ and $\sum_{i \leq n} S_i$.

Important bound: to guarantee $P_n(t) = 1$, we condition/-restrict so that $\sum_{n<\infty} \lambda_n^{-1} = \infty$ (LHS is expected time before population is infinity, or exceed any natural number).

By semantics, in contrast to that, $1 - \sum P_n$ is probability that $X(t) = \infty$. We must assume that "in finite time, no chance population go infinite"; if λ_ks small enough population go to infinity very fast; check the beginning equations.

Solving gives $P_n(t) = \prod_{0 \leq i \leq n-1} \lambda_i$ times a lot of stuff, with $n+1$ terms of $B_{i,n} \cdot e^{-\lambda_i t}$.

Yule Process. Prob of each member of pop give birth to one new member is $\beta h + o(h)$ in a time (interval) of length h. Assuming independence, $P(X(t+h) - X(t) = 1 | X(t) = n)$ is $n[\beta h + o(h)][1 - \text{earlier probability}]^{n-1}$, equal to $n\beta h + o_n(h)$.

The infinitesimal parameters are $\lambda_n = n\beta$ (directly proportional to pop size). This process is stochastic analogue of deterministic population model $dy/dx = \alpha y$.

Deterministic: rate (dy) proportional to pop size. Stoc: dy replaced by prob of unit increase (i.e. 1-increase) in dt. Computing gives $P_n = e^{-\beta t}(1 - e^{-\beta t})^{n-1}$ (can be computed by adding nCks in the previous P_n formula).

9 Pure-Death Process

Postulates. Complementing the "increasing Pure-Birth Proc" is the decreasing Pure-Death Proc. Moves successively through states $N \to 1$ and ultimately absorbed in 0. Specified through death params $\mu_k > 0$, and sojourn time in k is expo dist with param μ_k, all of them independent.

- $P(X(t+h) = k-1 | X(t) = k) = \mu_k + o(h)$, $1 \leq k \leq N$,
- Equal is 1−above probability; and grows/increases with probability 0.

Common (convention) to assign $\mu_0 = 0$. We have $P_N(t) = e^{-\mu_N t}$ and $P_n(t)$ similar "shape" as birth process; BUT the indices "from 0 to n" is replaced with "from n to N".

Linear Death Proc (complements Yule Proc). Define T as time of extinction (formally $:= \min\{t \geq 0 : X(t) = 0\}$). $T \leq t$ iff $X(t) = 0$. Then we have $F_T(t) = P(T \leq t) = P(X(t) = 0) = P_0(t)$ [chain of reasoning], giving us the [actually simple expression] of $(1 - e^{-\alpha t})^N, t \geq 0$.

Can be viewed in 2 other ways:

- 1. N individuals, lifetime expo dist with param α. $X(t)$ is survivors at time t.
- Proof roughly: $P(S_N > t)$ equal to $P(\min\{\text{death time of all indiv} > t\}$.
- 2. Approach in transition rates: each individual (take individual 1) has constant death rate $P(t < \xi_1 < t+h | t < \xi_1)$ equals $1 - e^{-\alpha h} = \alpha h + o(h)$.

Cable Failure. Support high-altitude weather balloon, with design load 1000 kg, design lifetime 100 years. We have the log-log equation between mean life μ_T, load l (of a single cable) be $\log_{10} \mu_T = 2 - 40 \log_{10} l$.

The individual fiber lifetime follows the r.v. $P(T \leq t) = 1 - \exp\left(-\int_0^t K(l(s)) ds\right), t \geq 0$. The hazard rate $r(t) = K(l(t)) l(t)^\beta / A$ [power law breakdown]; here $\beta = 40, A = 100$.

The system failure occurs on W_N (wait time to Nth fiber failure). The sojourn times S_k has expo with params $\mu_k = k \cdot K(NL/k) = k(NL/k)^\beta / A$; this is assuming that the beginning load is NL, and initially 1 fiber carries L load.

Can be directly calculated using earlier formula (of pure death process) or estimated by $EW_N = AL^{-\beta} \sum_{1 \leq k \leq N} 1/k(k/N)^\beta$, which is equal to $AL^{-beta} \sum 1/N(k/N)^{\beta-1}$. The sum is approximately $1/\beta$ by integral approximation. Compare with avg of single fiber lifetime of $\mu_T = A/L^\beta$.

10 Birth-Death Proc: Continuous-Time Analogue of Random Walk

Postulates. Define transition probabilities to be stationary (for now) $P_{ij}(t) := P(X(t+s) = j | X(s) = i)$.

- $P_{i,i+1}(h) = \lambda_i h + o(h)$, when $h \downarrow 0, i \geq 0$; and $P_{i,i-1}(h) = \mu_i + o(h)$, when $h \downarrow 0, i \geq 1$.
- $P_{ii}(h)$ equals 1-(sum of two terms above).
- $P_{ij}(0) = \delta_{ij}$. [idk why this is here]
- $\mu_0 = 0$, and $\mu_i, \lambda_i > 0$ for non-μ_0 parameters.

The matrix $[(-\lambda_0, \lambda_0, 0, 0, \ldots)(\mu_1, -(\lambda_1 + \mu_1), \lambda_1, 0, \ldots), \ldots]$ and so on is infinitesimal generator. The parameters are infinsim birth/death rates.

Note that $P_{ij} \geq 0$, $\sum_j P_{ij}(t) \leq 1$ (as they are probabilities) and the Chapman-Kolmogorov $P_{ij}(t+s) = \sum_k P_{ik}(t) P_{kj}(s)$ holds. The latter is because $i \to j$ in $t+s$ time, $X(t)$ moves to some state k in time t, then must from $k \to j$ in (remaining) time s.

Also to obtain $P(X(t) = n)$, this is equal to $\sum_i (q_i P_{ij}(t))$ where $q_i = P(\text{initial state} = i)$.

Sojourn Times Calculation. With assumptions already defined, calculate the distribution of r.v. S_i — time at state i, until it $X(t)$ leaves i. Turns out $S_i \sim \text{Expo}(-(\lambda_i + \mu_i)t)$; it enters $i+1, i-1$ with probability $\lambda_i / (\lambda_i + \mu_i)$ and (1 - minus that) respectively, after it leaves.

This is a motion ANALOGOUS to a random walk, but trnasitions occur at random times instead of fixed periods. This is called the "minimal" process (??)

Note: exists unique process with transition probability function $P_{ij}(t)$ with $\sum_j P_{ij} \leq 1$ and Chapman-Kolmo being true iff [insert formula, lazy; ch.2 p.11]. This is finding set of "all possible (λ_i, μ_i).

The Differential Equations to Compute Pij. [Must the equations follow P_{ij} ="some expression of P_{xy} with j fixed?]

We have the equations

- $P_{0j}' = -\lambda_0 P_{0j} + \lambda_0 P_{1j}$, t camouflaged (as usual, $t \geq 0$).
- $P_{ij}' = \mu_i P_{i-1,j} - (\lambda_i + \mu_i) P_{ij} + \lambda_i P_{i+1,j}$; this holds for $i \geq 1$ ($i = 0$ is a special case that it cannot regress, so it's the above equation instead).

Derivation is through Chapman-Kolmo $P_{ij}(t+h) = \sum_k P_{ik}(h) + P_{kj}(t)$, with h small. This equals $P_{i,i-1}(h) P_{i-1,j}(t) +$ 2 other expr $+ \sum_{\text{not important}} P_{ik}(h) P_{kj}(t)$. Last sum is $o(h)$. Div by h, $h \downarrow 0$ gives these eqns.

This is the backward equations: we divide $(0, t+h)$ into $(0, h)$ and $(h, t+h)$. Forward equations fixes i and do "last (small) step analysis" instead.

Linear Growth with Immigration. $\lambda_n = \lambda n + a$, $\mu_n = \mu n$. a may be interpreted as infinsim rate of increase "due to external source (i.e. immigration)".

We have $P_{i0}'(t) = -a P_{i0}(t) + \mu P_{i1}(t)$ (second set of eqns also similarly derived, with the lambda term to have $+a$) [forward equations]. Multiply the jth equation by j then sum to have $EX(t) = M(t) = \sum j P_{ij}(t)$ to have nice properties; $M'(t) = a + (\lambda - \mu) M(t)$.

If $X(0) = M(0) = i$, solution is $M = at + i$ when $\lambda = \mu$ and $\frac{a}{\lambda - \mu}(e^{(\lambda - \mu)t} - 1) + i e^{(\lambda - \mu)t}$.

11 Birth and Death Process Limiting Behavior, when No Absorbing States

Basic Properties

- Can be proved that $\lim_{t \to \infty} P_{ij}(t) = \pi_j \geq 0$. Exist and indep of init state i.
- It may happen that $\pi_j = 0$ for all j. When all are (strictly) positive, and satisfy $\sum \pi_j = 1$, it forms prob dist, called limiting distribution of the process.
- This is also a stationary distribution that $\pi_j = \sum_i \pi P_{ij}(t)$ — if it starts at i with probability π, at any time t it will be in i with that probability.

Proof of this follows from the differential equations, together with the first assumption (of the itemize) that the limit exists.

Rough proof: pass limit $t \to \infty$ and lim of right side (the no-derivative-side) exists, it follows that derivatives exist. And then, since prob is converging to constant, the limit of derivatives is 0. It produces "new equations" such as

- $0 = -\lambda_0 \pi_0 + \mu_1 \pi_1$,
- $0 = \lambda_{j-1} \pi_{j-1} - (\lambda_j + \mu_j) \pi_j + \mu_{j+1} \pi_{j+1}$, for $j \geq 1$.

Solving blablа and defining $\theta_0 = 1, \theta_j = \frac{\lambda_0 \ldots \lambda_{j-1}}{\mu_1 \ldots \mu_j}$, gives $\pi_{j+1} = \theta_{j+1} \pi_0$.

Having $\sum \pi_j = 1$, and summing all gives $1 = (\sum \theta_k) \pi_0$. If sum is infinity, necessarily $\pi_0 = 0$ (and hence $\pi_j = 0$). Otherwise, get $\pi_j = \theta_j \pi_0$ equals to θ_j/sum of all the thetas.

Linear Growth with Immigration lim-dist computation. We determine limit dist for $\lambda < \mu$. We get $\theta_k = [(\alpha/\lambda) + k - 1] C[k] \cdot (\lambda/\mu)^k$. Using infinite binomial we get $\sum \theta_k = (1 - \lambda/\mu)^{-\alpha/\lambda}$.

Thus π_0 is the reciprocal of that (change the exponent from negative to positive), and $\pi_k = \theta_k \cdot \pi_0$, for $k > 1$.

12 B/D Proc with Absorbing States (Longest Section, pages 16-22)

B/D Proc with $\lambda_0 = 0$ arise freqeuently (and correspondingly important). Zero state is absorbing in this case. Example is B/D Without Immigration (stays 0 if ever 0).

Probability of Absorption into 0. An important question to answer: must we always get there in the first place? How often do we expect to go there starting from, say, state i? Note that this is not a priori [theoretically always go to 0 by "simple reasoning"] since it may wander among $\mathbb{N}$ with probability 1 or it may drift to ∞.

Denote probability from i to be u_i. By familiar first step analysis, obtain $u_i = (\lambda_i / (\lambda_i + \mu_i)) u_{i+1}$ plus (1 - that parameter) times u_{i-1}. Another is considering "embedded random walk" with p_i = the first constant defined in u_i's equation.

Rewriting with $v_i = u_{i+1} - u_i$ gives $v_i = (\mu_i / \lambda_i) \cdot v_{i-1}$; defining $\rho_0 = 1$ and $\rho_i = \frac{\mu_1 \cdots \mu_i}{\lambda_1 \cdots \lambda_i}$, gives $v_i = \rho_i v_0$ and expanding to u_i gives $u_{i+1} - u_i = \rho_i (u_1 - 1)$.

Further computation gives that if $\sum \rho_i = \infty$, u_1 (and consequently u_n) is 1. Otherwise, can be shown that $u_m \to 0$ as $m \to \infty$. Throwing $m \to \infty$ once again solves for u_1 and hence

$$u_m = \frac{\sum_{i \geq m} \rho_i}{1 + \sum_{i \geq 1} \rho_i}.$$

Mean Time Until Absorption (from a fixed state, say m). We assume that absorption is certain (i.e. sum of rho's is infinite). Note that we cannot reduce our problem to embedded random walk since ACTUAL TIME in each state is relevant to the calculation of mean absorption time.

Recalling mean wait time in state i is $(\lambda_i + \mu_i)^{-1}$, letting ω_i be mean abs time from i, we deduce (once again, by first step analysis):

$$\omega_i = \frac{1}{\lambda_i + \mu_i} + \frac{\lambda_i}{\lambda_i + \mu_i} \omega_{i+1} + \frac{\mu_i}{\lambda_i + \mu_i} \omega_{i-1}$$

Write $z_i := \omega_i - \omega_{i+1}$, after some processing, gives $z_m = \left(\sum_{i \leq m} (1/\lambda_i)(\rho_m / \rho_i)\right) + \rho_m z_0$. Expanding gives

$$\frac{1}{\rho_m}(\omega_m - \omega_{m+1}) = \sum_{i \leq m} \frac{1}{\lambda_i \rho_i} - \omega_1.$$

We now analyze that equation. Since it is probabilistically evident that $\omega_m < \omega_{m+1}$, if $\sum_{i<\infty}$ of the fraction is infinity, ω_i is necessarily infinity.

Otherwise, letting $m \to \infty$ gives $\omega_1 = \text{sum} - \text{LHS}$. It is more involved but possible to prove LHS $= 0$, so we get ω_1 equals the sum. In any case, we get

$$\omega_m = \omega_1 + \sum_{k=1}^{m-1} \rho_k \sum_{j \geq k+1} \frac{1}{\lambda_j \rho_j}$$

with the sums are "continuations" of the first term (note that ω_1 has no "loose ρs" at the front of its expression).

The provided example is just Regular B/D Proc W/o Immigration.

Substitution of $a = 0, i = m$ (given that initial population is m) gives $M(t) = me^{(\lambda-\mu)t}$, exhibiting exponential growth/decay when $\lambda > \mu$ or $\lambda < \mu$.

We examine extinction phenomenon now (absorption in 0 of B/D Proc). Direct calculation yields $\rho_i = (\mu/\lambda)^i$ and $\sum \rho_i$ is finite (geometric sum) when $\lambda > \mu$ and infinite otherwise.

Eventual extinction, due to the first theorem, is $(\mu/\lambda)^m$ when $\lambda > \mu$ and 1 otherwise. Interestingly, when $\lambda = \mu$, process is sure to vanish eventually yet mean remain constant at init population level. These similar situations frequently arise when stochastic elements are present.

When $m = 1$ (assuming extinction is certain; that is, when $\lambda \le \mu$), mean time to absorption is $\sum 1/(\lambda_i \mu_i)$ equals to $1/\lambda \ln(\mu/(\mu-\lambda))$, when $\mu > \lambda$, and ∞ when $\mu = \lambda$.

Additional remarks on page 21.

- Many natural populations exhibit "density-dependent behavior" wherein the (indiv) birth rates ↓ or death rates ↑ as population ↑.
- These changes are ascribed to factors like limited food, increased predation, crowding, limited nesting sites. We also introduce a notion of environmental carrying capacity of K, upper bound that pop size cannot exceed.
- Since all individuals have chance to die, with finite K, all populations eventually go extinct. So our measure of population fitness is mean time to extinction ("mean population lifetime").
- Simplest model, birth params is $\lambda_n := n\lambda$ when $n < K$ and 0 for $n \ge K$.

The mean time to pop extinction is ω_1 given by $1/\mu \sum_{i \le K} 1/i \cdot (\lambda/\mu)^{i-1}$.

The first factor (that influence mean time to pop extinction) is $1/\mu$; mean lifetime of an indiv (since μ is indiv death rate). Thus, sum represent mean "generations" (or lifespans), a dimensionless quantity that we denote by $M_g = \mu\omega_1$, equalling $\sum 1/i(\theta)^{i-1}$ where $\theta = \lambda/\mu$ (represents reproduction rate).

When $K > 100$ or more, we have accurate approximations of $M_g \sim 1/\theta \ln(1/(1-\theta))$ if $\theta < 1$, $0.57721157 + \ln K$ for $\theta = 1$ and $1/K(\theta^k/(\theta-1))$ for $\theta > 1$.

13 Finite-State Cts-Time MC

MC Re-definitions. Cts time MC is Markov Process on states $\{0, 1, 2, \cdots\}$. Assume transition prob are stationary; $P_{ij}(t) = P(X(t+s) = j | X(s) = i)$. In this section we only consider when $|S|$ is finite, labeled $\{0, 1, \cdots, N\}$.

The Markov property asserts that $P_{ij}(t)$ satisfies

- $P_{ij}(t) \ge 0$,
- $\sum_{j=0}^{N} P_{ij}(t) = 1,\ 0 \le i, j \le N$,
- $P_{ik}(s+t) = \sum_{0 \le j \le N} P_{ij}(s) P_{jk}(t)$ for $t, s \ge 0$ (CHAPMAN-KOLMO relation).

We postulate in addition that $\lim_{t\to 0+} P_{ij}(t) = \delta_{ij}$ (initial states).

If $\mathbf{P}(t) := \|P_{ij}(t)\|$ then 3rd property can be rewritten (in matrix notation) as $\mathbf{P}(t+s) = \mathbf{P}(s)\mathbf{P}(t)$, $t, s \ge 0$. In addition, 4th property asserts that $P(t)$ (as functions) are cts at $t = 0$; it is equivalent to $\mathbf{P}(0) = \mathbf{I}$.

Proving some properties of P and Sojourn Times. It follows by taking $t, s = t, h$ and $t, s = (t-h, h)$ gives continuity.

Furthermore, we assume that it is also infinitesimally differentiable; or $\lim_{h\to 0+} P_{ij}(h)/h = q_{ij}$ for $i \ne j$. The sum of rows equal 1 gives $(1 - P_{ii}(h))/h =: q_i$ existing, and the equality $q_i = \sum_{j \ne i} q_{ij}$.

A note on sojourn times: when in state i, sojourns for Expo(q_i) time. Then jumps to $j \ne i$ with probability $p_{ij} = q_{ij}/q_i$, etc. The sequence of visited states by process (denoted $\xi_0, \xi_1, \cdots$) is called Embedded Markov Chain.

Assuming the Infinitesimal Properties have been verified, we derive explicit $P_{ij}(t)$; in terms of $\mathbf{A}$:= infinisim matrix.

The limit relations may be expressed (concisely) in the form $\lim_{h\to 0+} (\mathbf{P}(h) - \mathbf{I})/h = \mathbf{A}$. This shows that $\mathbf{A}$ is matrix derivative of $\mathbf{P}$ at $t = 0$, or $\mathbf{A} = \mathbf{P}'(0)$.

With that statement, and the multiplicative property, we get $\mathbf{P}'(t) = \mathbf{P}'(0)P(t)$, and the first term is just $\mathbf{A}$. Note that $\mathbf{P}'(t)$ is matrix whose elements are $P'_{ij}(t) = dP_{ij}(t)/dt$. Standard methods yield

$$\mathbf{P}(t) = e^{\mathbf{A}t} = \mathbf{I} + \sum_n \frac{\mathbf{A}^n t^n}{n!}.$$

Example 1. 2-state MC. Honestly nothing too interesting, just that the multiplication of two of those infinisim matrix is a multiple of itself, allowing for easy calculation.

We have $\mathbf{P}(t) = [(1-\pi, \pi), (1-\pi, \pi)]$ plus $[(\pi, -\pi), (-(1-\pi), 1-\pi)]e^{-\tau t}$; $\pi := \alpha/(\alpha+\beta)$, $\tau = \alpha + \beta$.

Example 2. Irreducible (every state communicate).

Passing limit of $t \to 0+$ gives expression $\pi(A + I) = \pi$ [we already assume limit exists]. It further gives $\sum_{i \ne j} \pi_i q_{ij} = \pi_j q_j$.

That equation has mass-balance interpretation. $\pi_j q_j$ is long run rate M-Proc leaves j, and that rate must be equal to long run rate M-Proc enters j (from some other state $i \ne j$), which is equal to the sum.

14 CH3. Basics of Renewal Proc.

Began with study of stoc-systems whose evolution (through time (not the time used as waiting times for a P-proc, for example)) was interspersed with renewals/-regeneration times, when the process began "anew" in a "statistical" sense. Today, the subject is viewed as study of general fns of iid nonnegative r.v.-s representing the successive intervals between renewals (WTF is this man on). Applicable in a wide variety of probability models.

Definition. Renewal (counting) proc $\{N(t) \mid t \ge 0\}$ is nonnegative int-valued stoc-proc that registers successive occurences of an event during (time-)interval $(0, t]$, where times between consecutive events ARE {positive; independent; iid} r.v.-s. Let successive durations between events be $\{X_k\}_{k \ge 1}$ (representing lifetimes of some units successively placed into service) (in other words: X_i elapsed time from $(i-1)^{\text{th}}$ event until occurence of i^{th} event.

We write $F(x) := P(X_k \le x), k \ge 1$ for common prob-dist. Basic stipulation for renewal-proc is $F(0) = 0$, signifiying X_i positive r.v.-s. We refer to $W_n = \sum_{i \le n} X_i$, $n \ge 1$ ($W_0 = 0$ be a convention) as wait-time until occurence of n^{th} event.

Relation between $\{X_k\}$ "interoccurence times" and $\{N_t, t \ge 0\}$ renewal proc is that $N(t)$ =number of indices n, for which $0 < W_n \le t$. In common practice, (counting-proc) $\{N(t)\}$ and partial sum proc $\{W(n)\}$ are interchangably called the "renewal proc".

Prototypical renewal model involves successive replacements of light bulbs. It's natural to assume successive lifetimes are statistically independent, with probabilistically indentical characteristic that it has common dist; in this proc, $N(t)$ records # of light-bulb-replacements up to time t. (JUST DRILL THIS IN, DEEP INSIDE YOUR HEAD RN)

Renewal Function. The principal objective of renewal theory is to derive properties of certain r.v. associated with those two (that we mention so often) from knowledge of common dist F. For example, compute expected # of renewals $E[N(t)] = M(t)$ (renewal fn).

To this end, several (pertinent := relevant/applicable to a particular matter) relationships & formulas are worth recording. We have the prob-law $W_m = \sum_{i \le n} X_i$, rephrased into $P\{W_n \le x\} = F_n(x)$ where $F_1(x$ =: $F(x))$ is known or prescribed, AND then in accordance with convolution formula,

$$F_n(x) = \int_0^\infty F_{n-1}(x-y)dF(y) = \int_0^x F_{n-1}(x-y)dF(y).$$

Connecting link. The fundamental conn-link between $wait-time$ proc and $renewal-count$ process is observation that

$$N(t) \ge k \Longleftrightarrow W_{\le} t.$$

Redundantly rephrased, this eqn asserts that num of renewals $\le$ time t is $\ge k$ iff kth renewal occured on or before time t. Since this equivalence is the basis for much that follows, reader should really verify instances of this (law) [This is further called (1.3) in this chapter].

The Actual Properties. $P\{N(t) \ge k\} = F_k(t)$, $t \in \mathbb{R}_{\ge 0}$, $k \in \mathbb{N}$. And consequently, $P\{N(t) = k\}$ equals to [geq k] - [geq $k+1$], so equals $F_k(t) - F_{k+1}(t)$.

For renewal function $M(t) = E[N(t)]$ we sum tail probabilities in the manner $E[N(t)] = \sum_{k \in \mathbb{N}} P\{N(t) \ge k\}$, and use first equality (i.e. the property after (1.3)) to get equal to $\sum_{k \in \mathbb{N}} F_k(t)$ (1.6).

Other r.v. of interest in renewal theory. Three are these: $\{\gamma_t := W_{N(t)+1} - t$ (excess/residual lifetime); $\delta_t := t - W_{N(t)}$ (current-life/age-r.v.); $\beta_t := \gamma_t + \delta_t$ (total life)$\}$.

An important identity enables us to eval $E[W_{N(t)+1}]$ in terms of the mean lifetime $\mu = E[X_1]$ of each unit and renewal-fn $M(t)$. Namely, it is true for every ren-proc that $E[W_{N(t)+1}] = E[X_1 + \ldots + X_{N(t)+1} = E[X_1] \cdot E[N(t) + 1]$; or $E[W_{N(t)+1}] = \mu \cdot (M(t) + 1)$ (1.7).

At first glance, this identity resemble formula for mean of random sum (that asserts that $E[X_1 + \ldots + E_N] = EX_1 EN$ when N is int-valued r.v. that is indep of $X_1, X_2, \ldots$; this doesn't apply here since the crucial diff. being "random num of summands $N(t) + 1$" not indep of "summands" themselves. (Show a more detailed result in Sec 3.3 later). For this reason, it isn't correct that $E[W_{N(t)}]$ be eval as prod of EX_1 and $EN(t)$ (identity becomes more intriguing and remarkable).

Proof. Fundamental equivalence (1.3) that sets $k := j - 1$ and t unchanged; expressed as indicator-r-v

$$\mathbf{I}\{N(t) \ge j-1\} = \mathbf{I}\{X_1 + \ldots + X_{j-1} \le t\}.$$

Since this indic-r-v a fn only of r.v. $X_1, \cdots, X_{j-1}$, it's indep of X_j, thus may eval

$$\begin{aligned} E\left[X_j \mathbf{I}\{\sum_{i \le j-1} X_i \le t\}\right] &= EX_j E\mathbf{I}\{\text{blah}\} \\ &= EX_j P\{\text{blah}\}(1.8) \\ &= \mu F_{j-1}(t). \end{aligned}$$

We calc $E[W_{N(t)+1}]$ which is sum of every X_i, $i \le N(t) + 1$. Further equals

$$= EX_1 + E\left[\sum_{j=2}^{N(t)+1} X_j\right]$$

$$= \mu + E\left[\sum_{j=2}^{\infty} X_j \mathbf{I}\{N(t) + 1 \ge j\}\right] \quad = \mu + \sum_{j=2}^{\infty} E\left[X_j \mathbf{I}\{X_1 + \cdots + X_{j-1} \ge j\}\right]$$

before further using lin-of-expect and pull sigma out, and formulating "$N(t) + 1 \ge j$" and applying the previous result gives it being equal to $\mu + \mu \sum_{j \ge 2}^{\infty} F_{j-1}(t)$ [using (1.8)] and equal to our desired result in (1.7) by a final application of (1.6).

15 Some Examples of Renewal Processes

Stochastic models often contain rand-times at which they/some-part-of-them, begin afresh (in a statistical sense). These renewal instants form natural embedded ren-proc-s, and they are found in many diverse fields of applied probability, including: {branching proc, insurance risk models, phenomena of pop-growth, evolutionary genetic mechanisms, engineering systems, econometric structures.} When a ren-proc is discovered embedded within a model, powerful results of renewal theory become available for deducing implications.

Brief Sketches of Renewal Situations. Several of these examples will be studied in more detail in later sections.

(a) P-Proc-s. $N(t)$, param λ, renewal cnt-ing proc having expo inter-occurence-dist $F(x) = 1 - e^{-\lambda x}$, $x \ge 0$. (this particular ren-proc possesses a host of special features, highlighted later in 3.3)

(b) Counter Proc-s. The times between successive electrical impulses/signals impinging on recording-device (counter) often assumed to form renewal proc.

Most physically realizable counters lock for some duration immed upon registering impulse; will not record impulses arriving during this dead period (only recorded when counter is free/unlocked). Should be emphasized that ren-proc of recorded impulses is secondary ren-proc derived from original ren-proc (that comprised totality of all arriving impulses).

(c) Traffic Flow. Distances between successive cars on an indefinitely long single-lane highway often assumed form ren-proc (so also are time-durations between consecutive-car passing a fixed loc).

(d) Ren-Proc Associated with Queues. In a single-server queueing proc there are (many) embedded natural ren-proc-s. We cite 2 examples: {(i) if customer arrival times form a renewal proc, then times of starts of successive busy periods generate a second ren-proc; (ii) for situation in which input proc (arrival pattern of customers) is Poisson, the succ-moments in which server passes from "a busy" to "a free" state determines a ren-proc.}

(e) Inventory System. In the analysis of most inventory-proc it is customary to assume that the "pattern demands" forms a ren-proc. Most of standard inventory-policies induce renewal sequences (e.g. times of replenishment of stock).

(f) Ren-Proc-s in MC-s. Let $Z_0, Z_1, \ldots$ be a recurrent MC. Suppose $Z_0 = i$, and consider the durations (elapsed num-of-steps) between successive visits to state i. Specifically, let $\{W_0 = 0;\ W_1 = \min\{n > 0; Z_n = i\};\ W_{k+1} = \min\{n > W_k;\ Z_n = i\}\ \}$.

Since each of these times is computed from same start-state i, the Markov property guarantees that $X_k = W_k - W_{k-1}$ are iid, and thus $\{X_k\}$ generates a ren-proc.

Block Replacement. Consider light bulb, whose life, measured in discrete units, is r.v. X where $P\{X = k\} = p_k$ for $k \geq 1$. Assuming one starts with fresh bulb and each is replaced by a new one when it burns out, let $M(n) = E[N(n)]$ expected replacements up to time n.

Because of scaling, in large building (office/factory) often cheaper (per-bulb basis) to replace everything, failed or not, than to replace a single bulb. "Block-replacement/Group-relamping" policy aims to take advantage of this. Strategy: replace bulb fail in $1, 2, \cdots, K-1$ [as they fail]. AND replace everything (fail or not) at K.

If c_1 is per-bulb block replacement cost, c_2 per bulb failure replacement cost ($c_1 < c_2$), mean total cost during the (whole) cycle is $c_1 + c_2 \cdot M(K-1)$, where $M(K-1) = E[N(K-1)]$ is mean num of failure-replacements {this is just considering cost for 1 bulb}. Since it consists of K periods, mean tot-cost per-bulb per-unit-time is $\theta(K)$ =above quantity/K. Want to choose block period $K = K^*$ that minimize cost rate $\theta(K)$.

The renewal func $M(n)$, or expected-num of replacements $\leq$ time n, solves equation $M(n) = F_x(n) + \sum_{k \leq n-1} p_k M(n-k)$ for $n \geq 1$. To derive: condition of (one) bulb life. If fails after time n, no additional replacement(s) (so, still have the $F_x(n)$). If at $k < n$ fails, then we have its failure, plus, on the average, $M(n-k)$ additional replacements during interval $[k+1, \cdots, n]$. Using LOTP to summ these contributions, we obtain

$$
\begin{aligned}
M(n) &= \sum_{k \geq n+1} p_k(0) + \sum_{k=1}^{n} p_k[1 + M(n-k)] \\
&= F_X(n) + \sum_{k \leq n} p_k M(n-k) \quad \text{[because } M(0) = 0\text{]},
\end{aligned}
$$

as asserted. There was a numerical example here, after the explanation (with maximum bulb life == 4).

We wish to elicit 1 more insight from this example. Forgetting block replacement, we continue to calc $M(5)$ until $M(10)$ when max-bulb-life is 4. Then $M(n) = M(n-1) + u_n$, probability that replacement occurs in period n. Then, we see that $u_n = M(n) - M(n-1)$ has limit $1/EX_1$. This makes sense, because if a light bulb lasts, on average, EX_1 time units, then the prob that it will need to be replaced in any period should approximate $1/EX_1$.

Actually, the relationship is not as simple as just stated; will be discussed further in 3.4 and 3.6.

16 P-Proc Viewed-as Renewal Proc

We mentioned earlier that inter-occurence times of this proc is $F(x) = 1 - e^{-\lambda x}$.

Renewal Fn: $N(t) \sim Pois(\lambda t)$, so $M(t) = E[N(t)] = \lambda t$.

Excess-Life: if $\lambda_t = t - W_{N(t)}$; observe that excess life at $t > x$ iff there are no renewals in interval $(t, t+x]$. This event has same probability as no renewals in $(0, x]$ (since P-proc has stationary independent increments).

In formal terms, we have $P\{\gamma_t > x\}$ equals $P\{N(x) = 0\} = e^{-\lambda x}$. So, excess life possesses same expo-dist $P\{\gamma_t \leq x\} = 1 - e^{-\lambda x}$ as ever life. This is another manifestation of memoryless property of expo-dist.

Current-Life: We know $\delta_t < t$ (cannot exceed t). For $x < t$, the current life (at t) exceeds x ifif no renewals in $(t-x, t]$, which AGAIN has prob $e^{-\lambda x}$.

Thus, current life follows truncated-expo-dist $P\{\gamma_t \leq x\} = [(1 - e^{-\lambda x}$ for $x < t$, 1 for $x \geq t]$.

Mean Total Life: using evaluation $EX = \int_0^\infty P(X > x)dx$ for mean of nonnegative rv, we have

$$
E\beta_t = E\gamma_t + E\delta_t = \frac{1}{\lambda} \int_0^t P\{\delta_t > x\}dx
$$

which is equal to $\frac{1}{\lambda} + \frac{1}{\lambda}(1 - e^{-\lambda t})$.

Observe that mean total life is SIGNIFICANTLY LARGER than mean life $1/\lambda = EX_k$ (of any particular ren-interval). A more striking is revealed when t large (process has been in operation for (a) long duration). Then mean total life $E\beta_t$ is approximately TWICE the mean life.

These facts appear at first paradoxical; re-examining defn of total life or β_t (intuitive explanation of seeming discrepancy). First, arbitrary time t is fixed; and such a procedure (of β_t measuring length of renewal interval containing point t) tend to favor (i.e. choose with higher likelihood) a lengthy renewal interval (than a shorter duration one). The phenomenon is known as length-biased-sampling, and occurs, well disguised, in a number of sampling situations.

Joint dist of γ_t, δ_t: event $\{\gamma_t > x, \delta_t > y\}$ occurs iff there are no renewals in the interval $(t-y, t+x]$, which has probability $e^{-\lambda(x+y)}$. Thus the prob of a realization of (x, y) [with $(x, y > 0)$] is $[e^{-\lambda(x+y)}$ if $y < t$, 0 if $y \geq t]$. Observe then that γ_t, δ_t independent, since joint dist factors as prod of their marginal dist-s.

17 Asymptotic Behavior of Ren-Proc-s

Many simple formulas that describe asymptotic behavior for large values of t of a general renewal process.

Elementary Renewal Theorem. The P-proc only renewal proc (in cts time) whose renewal fn $M(t) = E[N(t)]$ is exactly linear. ALL renewal fns are asymptotically linear, in the sense that

$$
\lim_{t\to\infty} \frac{M(t)}{t} = \lim_{t\to\infty} \frac{E[N(t)]}{t} = \frac{1}{\mu} (4.1)
$$

where $\mu = EX_k$ is mean inter-occurence time. This fundamental result (known as [the underlined "header"]) is invoked repeatedly to compute functionals during long-run-behavior of stoc-models having renewal-proc-s associated (with them). Note: (4.1) holds even when inter-occurence times have infinite mean, so RHS is then 0.

This theorem is so intuitively plausible that it has often been viewed as obvious; LHS (the most left-side) describes long-run mean-number of renewals per unit time; and RHS, is reciprocal of mean life of a component. (Obvious that) If component lasts on average μ time units, in the long-run these components will be replaced at rate of $1/\mu$ per unit time.

Example: Age Replacement Policies. Let $\{X_i\}$ iid positive, with finite $\mu = EX_k$. They are lifetime of items that are successively placed (in-service), next item commencing service immed-following failure of previous one, ERT tells us to expect to replace items (over the long run) at mean rate $1/\mu$ per-unit-time.

An age-repl-pol calls for repl item upon its failure/its reaching age T (whichever-occurs-first). We would expect that long-fraction of failure-repl (items fail before age T) will be $F(T)$ (and correspondingly, frac of planned-repl will be $1 - F(T)$). Renewal policy is $F_T(x) = F(x)$ for $x < T$, 1 for $x \geq T$, AND mean-renewal-duration is $\mu_T = \int_0^\infty [1 - F_T(x)]dx = \int_0^T [1 - F(x)]dx < \mu$.

ERT indicates long-run mean-repl is increased to $1/\mu_T$.

Now let $\{Y_i\}$ denote times between ACTUAL-SUCCESSIVE-failures; i.e. r.v. Y_1 is composed of "random num of time periods of length (exactly) T" (corresponding to repl-s not associated with failures) PLUS "a last time-period in which the dist is that a failure conditioned on failure-before-age-T" (that is, Y_1 has dist of $NT + Z$ where $P(N \geq k) = [1 - F(T)]^k$ and $P(Z \leq z) = F(z)/F(T)$, $0 \leq z \leq T$).

Hence

$$
\begin{aligned}
EY_1 &= \left(T[1 - F(T)] + \int_0^T [F(T) - F(x)]dx\right) / F(T) \\
&= \left(\int_0^T [1 - F(x)]dx\right) / F(T) = \mu_T / F(T).
\end{aligned}
$$

Sequence of $\{Y_i\}$ generates a renew-proc whose mean-rate of failures (in long-run) is $1/EY_1$ (inference relies on ERT). Depending on F, modified fail-rate $1/EY_1$ may possibly yield a lower fail-rate than $1/\mu$, rate where repl-s are made only upon failure.

Let's suppose each repl (whether planned/not), costs \$$K$, and each failure incurs an ADDITIONAL penalty of \$$c$. Multiplying these costs by the appropriate rates gives long-run-mean-cost-per-unit-time as fn of repl-age T:

$$
C(T) = \frac{K}{\mu_T} + \frac{c}{EY_1} = \frac{K + cF(T)}{\int_0^T [1 - F(x)]dx}.
$$

Renewal-Theorem for Cts-Lifetimes. It's tempting to conclude from ERT that $M(t)$ behaves like t/μ as $t \to \infty$, BUT the precise meaning of phrase "behaves like" is rather subtle. Example: if all lifetimes are 1 ($X_k = 1, \forall k \geq 1$) then $M(t) = \lfloor t \rfloor$. Since $\mu = 1$, $M(t) - t/\mu = \lfloor t \rfloor = t$, a fn that oscillates indefinitely between 0 and -1. While $M(t)/t = \lfloor t \rfloor / t \to 1 = 1/\mu$, it isn't clear in what sense that $M(t) \sim t/\mu$.

The fact is that if we "rule out" the periodic behvaior (??? this means we throw out periodic-looking processes (???)), then $M(t)$ behaves like t/μ in the sense described by the following theorem:

Let $M(t, t+h] = M(t+h) - M(t)$ denote mean num of renewals in interval $(t, t+h]$. The ren-thm states when periodic behavior is precluded, then

$$
\lim_{t\to\infty} M(t, t+h] = h/\mu \quad h > 0. \quad (4.2)
$$

In words: asymptotically, mean num of ren-s in an interval is proportional to interval's length, with prop-ity constant $1/\mu$.

A simple and prevalent situation in which ren-thm (4.2) is valid OCCURS WHEN lifetimes $X_1, X_2, \ldots$ are cts r.v. having density $f(x)$. In this case, ren-fn is differentiable, and (see (1.6): (1.6) states that $M(t) = \sum_k P(W_k \leq t) = \sum_k F_k(t)$)

$$
m(t) = \frac{dM(t)}{dt} = \sum_{n \geq 1} f_n(t), \quad (4.3)
$$

where $f_n(t)$ is density for $W_n = \sum_{i \leq n} X_i$. Now, (4.2) can be rewritten in the form $\dfrac{M(t+h) - M(t)}{h} \to \dfrac{1}{\mu}$ as $t \to \infty$, which, when h small, suggests that

$$
\lim_{t\to\infty} m(t) = \lim \frac{dM(t)}{dt} = \frac{1}{\mu}.
$$

IF IN ADDITION to being cts, the lifetimes $X_1, X_2, \ldots$ HAVE FINITE MEAN AND VARIANCE (say, variance is σ^2), then ren-thm can be refined to include a 2nd term. Under stated conditions, we have

$$
\lim_{t\to\infty} \left(M(t) - \frac{t}{\mu}\right) = \frac{\sigma^2 - \mu^2}{2\mu^2}.
$$

There was an (easy) example with $f(x) = xe^{-x}$ here.

Asymptotic Dist of N(t). We have $\lim_{t\to\infty} EN(t)/t = 1/\mu$, implies asymptotic mean of $N(t)$ is approx t/μ. When $\mu = EX_k$, $\sigma^2 = \mathrm{Var}X_k$ are finite, the asymptotic var of $N(t)$ is $\lim_{t\to\infty} \mathrm{Var}N(t)/t = \sigma^2/\mu^3$. That is, asymptotic var of $N(t)$ is approx $t\sigma^2/\mu^3$.

We get following convergence to (standard) normal dist

$$
\lim_{t\to\infty} \left(\frac{N(t) - t/\mu}{\sqrt{t\sigma^2/\mu^3}} \leq x\right) = \Phi(x).
$$

In words, for large values of t, num of renewals $N(t)$ is approx-ly normal-ly dist with (mean/t) $1/\mu$ and (var/t) σ^2/μ^3.

Limiting Dist of Age and Excess Life. We assume lifetimes $X_1, \ldots$ are cts r.v. with finite mean μ. Let $\gamma_t = W_{N(t)+1} - t$ be excess life at t. It has limiting dist

$$
\lim_{t\to\infty} P(\gamma_t \leq x) = \frac{1}{\mu} \int_0^x [1 - F(y)]dy. \quad (4.9)
$$

The right side of (4.9) is a valid dist fn.

(VERIFICATION: by defn on 3.1, we have $P\{W_n \leq x\} = F_n(x) (= F_1(x) = F(x))$. So int_0^∞ of $P\{W_n \geq x\}$ is $E[W_n]$, and we have $1/\mu = 1/E[W_n]$, so it is verified!)

which we denote by $H(x)$. The corresponding density is $h(y) = \mu^{-1}[1 - F(y)]$. The mean $H(y)$ is $\int_0^\infty yh(y)dy$ which after exapnding $h(y)$ and regrouping it into

$$
\frac{1}{\mu} \int_0^\infty y \int_y^\infty f(t)dtdy
$$

and switching the order of integral calculations, it becomes $1/(2\mu) \int_0^\infty t^2 f(t)dt$, equalling $\frac{\sigma^2 + \mu^2}{2\mu}$, with σ^2 common var of the lifetimes.

There is a strange note:

$$\{\gamma_t \geq x, \delta_t \geq y\} \longleftrightarrow \{\gamma_{t-y} \geq x + y\}.$$

It follows that (after defining the limit for $P\{\gamma_t \geq x, \delta_t \geq y\}$)

$$\lim_{t\to\infty} P\{\delta_t \geq y\} = \mu^{-1} \int_y^\infty [1 - F(z)]dz$$

which is (finally) equal to $1 - H(y)$.

Proof of ERT. It is always the case that $t < W_{N(t)+1}$. By (1.7) we have $t < \mu[1 + M(t)]$, and therefore $M(t)/t > \mu^{-1} - t^{-1}$ (to get this, we divide the previous statement by $\mu \cdot t$). It follows from this that $\liminf_{t\to\infty} M(t)/t \geq \mu^{-1}$.

To establish opposite inequality, let $c > 0$ arbitrary ("truncate-point") and set $X_i^c = [X_i \text{ if } X_i \leq c, c \text{ if } X_i > c]$. And consider renewal proc having lifetimes $\{X_i^c\}$. Let $W_n^c, N^c(t)$ denote wait-times and cnt-proc, respectively, for this "truncated-ren-proc" generated by $\{X_i^c\}$.

Since r.v. X_i^c unif-bounded by c, it's clear that $t + c \geq W_{N^c(t)+1}$, and therefore

$$t + c \geq E(\text{above RHS}) = \mu^c[1 + M^c(t)];$$

where (here we describe explicitly what each term means) $\mu^c = EX_1^c = \int_0^c [1 - F(x)]dx$, AND $M^c(t) = EN^c(t)$.

Obviously $X_i^c \leq X_i$ entails/implies $N^c(t) \geq N(t)$; and so $M^c(t) \geq M(t)$. So, we have $t + c \geq \mu^c[1 + M(t)]$, and by rearrangement (and dividing by $\mu_c \cdot t$), we have

$$M(t)/t \leq \mu_c^{-1} + t^{-1}(c/\mu^c - 1)$$

(yes, he changes notation from mu^c into μ_c to fit $^{-1}$ into μ_c^{-1}). Hence, $\limsup_{t\to\infty} M(t)/t \leq \mu_c^{-1}$ [we pass $t \to infty$ and so we have t^{-1} vanishes to 0, disabling whatever positive/negative terms we have multiplied by t^{-1}. If we need c to be large we can set t to be LARGER here in this limit].

Since $\lim_{c\to\infty} \mu^c = \lim_{c\to\infty} \int_0^c [1-F(x)]dx$ and pass $c \to \infty$ replacing c in integral-upper-bound to ∞; and this quantity is equal to μ. Since we have LHS is fixed, we have $\limsup_{t\to\infty} M(t)/t \leq \mu^{-1}$.

The two bounds imply the assertion.

18 Generalizations + Variations on Renewal Proc-s

Delayed Ren-Proc-s. We continue to assume $\{X_k\}$ are all indep positive r.v., but only $X_2, X_3, \ldots$ iid with dist fn F, while X_1 has possible a different dist fn G. Such process called delayed ren-proc.

Such processes will arise when component that is in operation (i.e. running) at time $t = 0$ is not new, BUT all subsequent replacements are new. (Example: suppose time in origin taken y time units after start of an ordinary ren-proc; then time to first renewal will have DISTRIBUTION OF EXCESS LIFE ("remaining life") at time y of an ordinary renewal process.

Let $W_0 = 0$ and $W_n = \sum_{i\leq n} X_i$. We have (to distinguish mean num of renewals in delayed process and renewal function associated with distribution F) $M_D(t) = E[N(t)]$ and $M(t) = \sum_{k\in\mathbb{N}} F_k(t)$.

For delayed process, elementary renewal thm is $\lim_{t\to\infty} M_D(t)/t = 1/\mu$ with $\mu = E[X_2]$ and renewal theorem states that

$$\lim_{t\to\infty}[M_D(t) - M_D(t-h)] = h/,$$

provided $X_2, X_3, \ldots$ are cts rv.

Stationary Ren-Proc-s. Delayed ren-proc for which first life has dist fn

$$G(x) := \mu^{-1} \int_0^x [1 - F(y)]dy$$

is called stationary-ren-proc. We are attempting to model a ren-proc that began indefinitely far in the past. So that remaining life (excess life) of item in service (at time origin) has limiting distribution of excess life in an ordinary ren-proc.

We recognize G as this limiting distribution.

It is anticipated that such a proc exhibits a number of stationary, or TIME-INVARIANT properties. For a stationary ren-proc, we have

$$M_D(t) = E[N(t)] = t/\mu, \text{and } P(\gamma_T^D \leq x) = G(x)$$

for all t. Thus, what is IN GENERAL ONLY AN ASYMPTOTIC renewal-relation becomes an IDENTITY (or an CHARACTERISTIC-IDENTIFIER) that holds for all t (in a stationary ren-proc).

Cumulative + Related Ren-Proc-s. Suppose that associated with the i^{th} unit, or lifetime interval, is a 2nd r.v. Y_i (with $\{Y_i\}$ identically distributed) IN ADDITION to lifetime X_i. We allow X_i, Y_i to be dependent (for every i) but assume that pairs

$$(X_1, Y_1), (X_2, Y_2), \ldots \text{ are pairwise independent.}$$

We use notation $F(x) = P(X_i \leq x)$ and $G(y) = P(Y_i \leq y)$, $\mu = E[X_i]$ and $\nu = E[Y_i]$.

A number of problems (of practical and theoretical interest) have a natural formulation in those terms.

Mini-Theorem/Mini-Section 1. Ren-proc-s involving to components to each ren-interval.

Suppose Y_i represent a portion of duration X_i. We depict in Figure 5.1. that the Y portion occurs at beginning of the interval, BUT this assumption is not essential. Let $p(t)$ be probability that t falls in a Y portion of some ren-interval. When $X_1, X_2, \ldots$ are cts rv, ren-thm implies the following (important) asymptotic eval:

$$\lim_{t\to\infty} p(t) = E[Y_1]/E[X_1].$$

Some concrete examples:

Replacement Model. Consider replacement model (in) whcih replacement is not instantaneous. Let Y_i be operating time, Z_i be lag period preceding installment of $(i+1)^{\text{st}}$ operating unit (upon breakdown/renewal after i^{th} unit breaks). We assume that sequence of times between successive replacements $X_k = Y_k + Z_k$ $(k \in \mathbb{N})$ constitutes ren-proc. Then, $p(t) \to EY_1/EX_1$.

Queueing Model. Some desigmated place where a service of some kind is being rendered. It is assumed that "time-between-arrivals", or "inter-arrival time" and "time-spent-in-providing-service" for a given customer are governed by probabilistic laws.

If arrivals to queue follow P-Proc intensity λ, THEN successive times X_k from commencement of k^{th} busy period to start of next busy period form a ren-proc. (Busy period := uninterrupted-duration when queue is-not-empty). Each X_k composed of a busy portion Y_k and idle portion Z_k, then, $p(t) \to EZ_1/EX_1$.

Peter Principle. Asserts that worker will be promoted UNTIL finally reaching a position in which he/she is incompetent. When happens, person stays in that job until retirement.

Consider following model: person selected at random (from the population) and placed in job. Person competent → he/she remains at the job for random time, having cdf F and mean μ AND is promoted. Person incompetent → remains for random time having cdf G and mean $\nu > \mu$ AND retires. Once job vacated, another person selected at random AND process repeats. Assume infinite population contains fraction p of competent people and $q = 1 - p$ incompetent ones.

In long run, what fraction of position time held by incompetent person? Renewal occurs every time position is filled, therefore mean duration of ren-cycle is

$$E[X_i] = p\mu + (1-p)\nu.$$

To answer qn, let $Y_k = X_k$ if k^{th} person is incompetent, and $Y_k = 0$ if k^{th} person is competent. Then, long-run-fraction of time that position is held by incompetent person is

$$\frac{EY_1}{EX_1} = \frac{(1-p)\nu}{p\mu + (1-p)\nu}.$$

Mini-Theorem/Mini-Section 2. Cumulative Processes.

Interpret Y_i AS A COST/VALUE associated with i^{th} renewal cycle. A class of problems with a natural setting (in this general context of pairs (X_i, Y_i) where X_i generated a ren-proc, will now be considered. Interest here focuses on so-called cumulative-proc

$$W(t) = \sum_{k=1}^{N(t)+1} Y_k,$$

the accumulated costs/value up to time t (assuming transactions are made at THE BEGINNING OF RENEWAL CYCLE).

ERT asserts in this case that

$$\lim_{t\to\infty} E[W(t)]/t = EY_1/\mu.$$

This eqn justifies the interpretation of $E[Y_1]/\mu$ as a long-run-mean-cost/value per unit-time.

Replacement Models. Age-Repl-Pol (see Section 3.3 and 3.4 and example titled "Age-Replacement-Policies") states that planned replacement at age T costs c_1 dollars, while failure replaced at time $x < T$ costs c_2 dollars.

If Y_k is cost incurred at the k^{th} replacement cycle, then $Y_k = [c_1$ with probability $1-F(T)$, c_2 with probability $F(T)]$. AND $E[Y_k] = c_1[1 - F(T)] + c_2F(T)$. Since expected length of replacement-cycle is

$$E[\min\{X_k, T\}] = \int_0^T [1 - F(x)]dx$$

(the T at the second argument passed to min() fn corresponds to upper-bound of integral from 0 to "an indefinite number" (default = ∞)), we have long-run-cost per unit-time is

$$\frac{c_1[1 - F(T)] + c_2F(T)}{\int_0^T [1 - F(x)]dx}$$

and in any particular situation, a routine calculus/recourse to numerical-computation produces value of T that minimizes long-run-cost per unit-time.

Under block-repl-policy, there is one planned-replacement every T units of time, and on average, $M(T)$ failure replacements. So, expected cost is $E[Y_k] = c_1 + c_2M(T)$ (and long-run-mean-cost per unit-time is [that value/T]).

Risk Theory. Suppose claims arrive at an insurance company according to ren-proc with inter-occurence times $X_1, X_2, \ldots$. Let Y_k be magnitude of k^{th} claim. Then, $W(t) = \sum_{k=1}^{N(t)+1} Y_k$ represents cumulative-amount of time claimed up to (time) t, AND the long-run-mean-claim rate is

$$\lim_{t\to\infty} E[W(t)]/t = EY_1/EX_1.$$

19 CH4: Everything About Brownian Motion.

Brownian motion and Gaussian proc arose naturally in 20th century as an attempt to explain ceaseless irregular motionos of tiny particles suspended in a liquid (like dust mots floating in air). Now has many applications.

BM Stoc-Proc. Cts-time, cts-state-space M-Proc. $B(t)$ is y component (as function of time) of particle, and x be the position of particle at t_0; $B(t_0) = x$.

Let $p(y, t \mid x)$ be pdf, in y, of $B(t_0 + t)$ [there are 2 dependent var and 1 independent var]; given $B(t_0) = x$ (any t_0 gives the same pdf-value; i.e. transitions are stationary).

Properties: $\{p(y, t \mid x) \geq 0$; $\int_{-\infty}^\infty p dy = 1$; $\lim_{t\to 0} p = 0$ for $y \neq x$; Diffusion-Equation: $\frac{\partial p}{\partial t} = \frac{1}{2}\sigma^2 \frac{\partial^2 p}{\partial x^2}$ (σ^2 diffusion coefficient)$\}$.

Choosing proper scale (so $\sigma^2 = 1$) can verify

$$p = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{1}{2t}(y-x)^2\right)$$

is solution of diffusion-eqn; and the p solution as normal pdf (mean x, var t).

Use $\phi(z) = \sqrt{1/(2\pi)} \exp(-1/2z^2)$ and $\Phi(z) = \int_{-\infty}^z \phi(x)dx$. The pdf (at z) of normal dist with var t is $1/\sqrt{t}\phi(z/\sqrt{t})$ and cdf (at z) is $\Phi(z/sqrtt)$. In this notation, transition density of p is $\phi_t(y - x)$ [before "scaling" down to std-normal] and $P(B(t) \leq y \mid B(0) = x) = \Phi((y-x)/sqrtt)$.

Properties of BM. Every increment is normal dist mean 0, var $\sigma^2 t$ (σ^2 fixed param), disjoint intervals are indep-r.v.-s, $B(0) = 0$ and $B(t)$ cts fn of t.

Remark. Look at Brownian-displacement after small elapsed time Δt. Std of Br-displacement is $\sqrt{\Delta t} >>> \Delta t$ (manifestation of erratic movements of Brownian particle). BUT the variance, being linear in time (and not the stdev) IS THE ONLY POSSIBILITY if displacements over disjoint-time-intervals are stationary and indep (can prove by "as if we are solving Olympiad-Cauchy-additive-functions").

Also, $\mathrm{Cov}[B(s), B(t)] = \sigma^2 \min(s, t)$ with $s, t > 0$.

CLT and Invariance Principle.

this is basically CLT with extra steps: modelling $S_n = \sum_{i \le n} \xi_i$ as sum of n iid r.v. ξ_i have 0 mean and 1 variance. In this case, CLT asserts

$$\lim_{n\to\infty} P(S_n/\sqrt{n} \le x) = \Phi(x).$$

[Similarly], Functionals computed for BM can often serve as excellent approximations [for BM]. Especially, for analogous/stationary partial-sum processes; as we now explain.

As a function in cts var t, observe the r.v. $B_n(t) := S_{\lfloor nt \rfloor}/\sqrt{n}$. Observe $B_n(t)$ is $S_k/n = \sqrt{\lfloor nt \rfloor}/\sqrt{n} \cdot S_k/\sqrt{k}$. Because $S_k/\sqrt{k}$ has unit variance, var of $B_n(t) = \lfloor nt \rfloor / n$, converges to t as $n \to \infty$. $B_n(t)$ reasonably-behave like std Br-motion, at least when n is large.

This convergence of sequence of stoc-proc-s to std-Br-motion is termed invariance principle. Asserts some functionals of part-sum-proc iid distributed 0 mean, 1 variance shouldn't-depend-too-heavily on (i.e. should be invariant of) actual dist of summands (provided "summands not too badly behaved").

Example: summands have distribution $\xi = \pm 1$. This random walk reaches $-a$ before b with prob $b/(a+b)$, so std-Br-mot reaches $-a$ before b probability $b/(a+b)$.

Finally, invoking the invar-principle for 2nd (and more general) time, the evaluation should hold approximately for arbitrary partial sum proc (provided iid, 0 mean, 1 var).

Gaussian Process. Random vector $X_1, \ldots, X_n$ said to have multivar-norm-dist or joint-norm-dist if every lincombin $\sum \alpha_i X_i$ ($\alpha_i \in \mathbb{R}$) is univariate-norm-dist.

This multivar-norm-dist specified by two (sets of) params: $\{\mu_i = E[X_i],\ \Gamma_{ij} = \mathrm{Cov}(X_i, X_j)\}$.

Let T abstract set, $\{X(t)\}$ stoc-procl call $\{X(t)\}$ Gaussian-proc if for finite subset $\{t_i\}_{i \le n}$ of T, the random vector $(X(t_i))_{i \le n}$ has multivar-norm-dist.

Every Gauss-proc described uniquely by two params, $\mu(t) = E[X(t)], t \in T$ AND $\Gamma(s,t) = E[(X(s) - \mu(s))(X(t) - \mu(t))]$ with $s, t \in T$.

Note that cov-fn is pos-definite in sense that $\forall n \ge 1, \alpha_i \in \mathbb{R}, t_i \in T$, we have

$$\sum_{i \le n} \sum_{j \le n} \alpha_i \alpha_j \Gamma(t_i, t_j) \ge 0,$$

which can be proved via

$$E[\sum_{i \le n} \alpha_i \left(X(t_i) - \mu(t_i)\right)]^2 \ge 0$$