Department of Mathematics, IST — Probability and Statistics Unit

Introduction to Stochastic Processes

1st. Test	2nd. Semester — 2012/13
Duration: 1h30m	2013/04/19 — 5PM, Room P8

- Please justify all your answers.
- This test has TWO PAGES and THREE GROUPS. The total of points is 20.0.

Group 0 — Introduction to Stochastic Processes **2.5 points**

A Bernoulli process with parameter $p = \frac{1}{2}$ has already been used in the investigation of geomagnetic reversals,[1] *with Bernoulli trials separated by 282 ky (i.e., 282 thousand years).*

(a) *Consider the stochastic process $\{S_n : n \in \mathbb{N}\}$, where S_n represents the number of geomagnetic reversals in $n \times 282$ ky.*
Is this stochastic process (second order weakly) stationary? (1.0)

- **Stochastic process**
$\{X_i : i \in \mathbb{N}\} \overset{i.i.d.}{\sim} \text{Bernoulli}(p = \frac{1}{2})$
- **Another stochastic process**
$\{S_n = \sum_{i=1}^n X_i : n \in \mathbb{N}\}$
S_n = number of geomagnetic reversals in $n \times 282$ ky
$S_n \sim \text{Binomial}(n, p = \frac{1}{2})$
- **Investigating the 2nd. order weak stationarity**
On one hand, $E(S_n) = np$ depends on the time (n), thus, the stochastic process $\{S_n : n \in \mathbb{N}\}$ is not 1st. order weakly stationary. On the other hand, 2nd. order weak stationarity implies 1st. order weak stationarity. Consequently, this stochastic process is not 2nd. order weakly stationary.
- **[Obs.**
For $n, s \in \mathbb{N}$,

$$\begin{aligned} cov(S_n, S_{n+s}) &= cov(S_n, S_{n+s}) \\ &= cov\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i + \sum_{j=n+1}^{n+s} X_j\right) \\ &\overset{X_i\ i.i.d.}{=} cov\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i\right) + 0 \\ &= V(S_n) \\ &= np(1-p), \end{aligned}$$

which depends on the time (n), thus, $\{S_n : n \in \mathbb{N}\}$ is not 2nd. order weakly stationary.]

[1] *A geomagnetic reversal is a change in the orientation of Earth's magnetic field such that the positions of magnetic north and magnetic south become interchanged (http://en.wikipedia.org/wiki/Geomagnetic_reversal).*

(b) *Find the probability that the number of Bernoulli trials needed to observe 4 geomagnetic reversals does not exceed 10.* (1.5)

- **New r.v.**
T_k = number of Bernoulli trials needed to observe exactly k geomagnetic reversals
$T_k \sim \text{NegativeBinomial}(k, p)$ (se formulae!)
- **Requested probability**

$$\begin{aligned} P(T_k \le x) &= F_{NegativeBin(r,p)}(x) \\ &\overset{form.}{=} 1 - F_{Binomial(x,p)}(k-1) \\ &= 1 - F_{Binomial(10,1/2)}(4-1) \\ &\overset{tables}{=} 1 - 0.1719 \\ &= 0.8281. \end{aligned}$$

- **Obs.**
$P(T_4 \le 10) = P(S_{10} \ge 4)$.

Group 1 — Poisson Processes **9.5 points**

1. *Admit outline accesses from within a local phone network are governed by a Poisson process with rate $\lambda = 1$ access per minute.*

(a) *Find the joint probability that the cumulative number of accesses is equal to 2 at time 1 minute, 3 at time 2 minutes, and 5 at time 3 minutes.* (1.5)

- **Stochastic process**
$\{N(t) : t \ge 0\} \sim PP(\lambda = 1)$
$N(t)$ = cumulative number of outline accesses at time t
= number of outline accesses by time t
$N(t) \sim \text{Poisson}(\lambda t = t)$
- **Requested probability**

$$\begin{aligned} P[N(1) = 2, N(2) = 3, N(3) = 5] &= P[N(1) = 2, N(2) - N(1) = 3 - 2, \\ &\quad N(3) - N(2) = 5 - 3] \\ &\overset{indep.\ incr.}{=} P[N(1) = 2] \times P[N(2) - N(1) = 3 - 2] \\ &\quad \times P[N(3) - N(2) = 5 - 3] \\ &\overset{station.\ incr.}{=} P[N(1) = 2] \times P[N(2-1) = 3 - 2] \\ &\quad \times P[N(3-2) = 5 - 3] \\ &= P[N(1) = 1] \times \{P[N(1) = 2]\}^2 \\ &\overset{N(1) \sim Poisson(1)}{=} e^{-1}\frac{1^1}{1!} \times \left(e^{-1}\frac{1^2}{2!}\right)^2 \\ &= \frac{e^{-3}}{4} \\ &\simeq 0.012448. \end{aligned}$$

(b) *What is the probability that the cumulative number of accesses at time 2 minutes exceeds 10, given that the cumulative number accesses is equal to 20 at time 4 minutes?* (1.5)

- **R.v.**
  $N(s) \mid N(t) = n$, $0 < s < t$, $n \in \mathbb{N}$
  $(N(s) \mid N(t) = n) \sim \text{Binomial}(n, s/t)$ (see form. for NNPP)

- **Requested probability**

$$\begin{aligned} P[N(2) > 10 \mid N(4) = 20] &= 1 - P[N(2) \leq 10 \mid N(4) = 20] \\ &= 1 - F_{Binomial(20,2/4=0.5)}(10) \\ &\overset{tables}{=} 1 - 0.5881 \\ &= 0.4119. \end{aligned}$$

- **[Obs.**
  For $0 < s < t$, $n \in \mathbb{N}$, $x = 0, 1, \ldots, n$,

$$\begin{aligned} P[N(s) = x \mid N(t) = n] &= \frac{P[N(s) = x, N(t) = n]}{P[N(t) = n]} \\ &= \frac{P[N(s) = x, N(t) - N(s) = n - x]}{P[N(t) = n]} \\ &\overset{indep.\, incr.}{=} \frac{P[N(s) = x] \times P[N(t) - N(s) = n - x]}{P[N(t) = n]} \\ &\overset{station.\, incr.}{=} \frac{P[N(s) = x] \times P[N(t-s) = n - x]}{P[N(t) = n]} \\ &\overset{N(z) \sim Poisson(\lambda z)}{=} \frac{e^{-\lambda s} \frac{(\lambda s)^x}{x!} \times e^{-\lambda(t-s)} \frac{[\lambda(t-s)]^{n-x}}{(n-x)!}}{e^{-\lambda t} \frac{(\lambda t)^n}{n!}} \\ &= \binom{n}{x} (s/t)^x (1 - s/t)^{n-x}, \end{aligned}$$

  which is the p.f. of a Binomial$(n, s/t)$ r.v.]

2. *Harry owns a vegetarian food stand that is open from 8:00 to 17:00 and admits that the customers arrive to it according to a non-homogeneous Poisson process with time dependent rate equal to*

$$\lambda(t) = \begin{cases} 10, & 8 \leq t \leq 11 \\ 20, & 11 < t \leq 13 \\ 15, & 13 < t \leq 17. \end{cases}$$

(a) *Derive the associated mean value function and obtain the expected number of arrivals to the vegetarian food stand between 10:00 and 14:00.* (2.0)

- **Stochastic process**
  $\{N(t) : 8 \leq t \leq 17\} \sim NHPP(\lambda(t))$
  $N(t)$ = number of arrivals to the Harry's vegetarian food stand until time t

$$\lambda(t) = \text{intensity function} = \begin{cases} 10, & 8 \leq t \leq 11 \\ 20, & 11 < t \leq 13 \\ 15, & 13 < t \leq 17 \end{cases}$$

- **Mean value function**
  For $8 \leq t \leq 17$,

$$\begin{aligned} m(t) &\overset{form.}{=} \int_8^t \lambda(z)\, dz \\ &= \begin{cases} \int_8^t 10\, dz = 10(t-8), & 8 \leq t \leq 11 \\ m(11) + \int_{11}^t 20\, dz = 30 + 20(t-11), & 11 < t \leq 13 \\ m(13) + \int_{13}^t 15\, dz = 70 + 15(t-13), & 13 < t \leq 17 \end{cases} \end{aligned}$$

- **Requested expected value**

$$\begin{aligned} E[N(14) - N(10)] &\overset{form.}{=} m(14) - m(10) \\ &= [70 + 15(14-13)] - [10(10-8)] \\ &= 85 - 20 \\ &= 65. \end{aligned}$$

(b) *Compute the probability that the second customer arrives to the Harry's vegetarian food stand between 8:30 and 9:00.* (2.0)

- **R.v.**
  S_2 = time of the 2nd. arrival

- **Requested probability**

$$\begin{aligned} P(8.5 < S_2 \leq 9) &= P(S_2 \leq 9) - P(S_2 \leq 8.5) \\ &\overset{form.}{=} P[N(9) \geq 2] - P[N(8.5) \geq 2] \\ &= \{1 - P[N(9) \leq 1]\} - \{1 - P[N(8.5) \leq 1]\} \\ &= P[N(8.5) \leq 1] - P[N(9) \leq 1] \\ &\overset{N(t) \sim Poisson(m(t))}{=} F_{Poisson(m(8.5))}(1) - F_{Poisson(m(9))}(1) \\ &= F_{Poisson(10 \times (8.5-8))}(1) - F_{Poisson(10 \times (9-8))}(1) \\ &= F_{Poisson(5)}(1) - F_{Poisson(10)}(1) \\ &\overset{tables}{=} 0.0404 - 0.0005 \\ &= 0.0399. \end{aligned}$$

3. *Suppose that the number of requests to a web server follows a conditional Poisson process with random rate Λ (in requests per minute) and admit that $\Lambda \sim Gamma(\alpha, \beta)$, where $\alpha, \beta > 0$.*

   Derive expressions for the expected value, the variance and the moment generating function of the number of requests to the web server by time t ($t > 0$). (2.5)

- **Stochastic process**
  $\{N(t) : t \geq 0\} \sim Conditional PP(Gamma(\alpha, \beta))$, $\alpha, \beta > 0$
  $N(t)$ = number of requests to a web server until time t

- **Random arrival rate**
  $\Lambda \sim \text{Gamma}(\alpha, \beta)$
  $E(\Lambda) = \frac{\alpha}{\beta}$

$V(\Lambda) = \frac{\alpha}{\beta^2}$

- **Distribution of N(t) conditional to $\Lambda = \lambda$, etc.**

  $(N(t) \mid \Lambda = \lambda) \sim \text{Poisson}(\lambda t)$

  $E[N(t) \mid \Lambda = \lambda] = \lambda t$

  $V[N(t) \mid \Lambda = \lambda] = \lambda t$

- **Requested expected value**

$$\begin{aligned} E[N(t)] &= E\{E[N(t) \mid \Lambda]\} \\ &= E(\Lambda t) \\ &= \frac{\alpha}{\beta} \times t \end{aligned}$$

- **Requested variance**

$$\begin{aligned} V[N(t)] &= V\{E[N(t) \mid \Lambda]\} + E\{V[N(t) \mid \Lambda]\} \\ &= V(\Lambda t) + E(\Lambda t) \\ &= \frac{\alpha}{\beta^2} \times t^2 + \frac{\alpha}{\beta} \times t \\ &= \frac{\alpha t}{\beta} \times \left(\frac{t}{\beta} + 1\right) \end{aligned}$$

- **Requested m.g.f.**

$$\begin{aligned} E\left[e^{sN(t)}\right] &= E\left\{E[e^{sN(t)} \mid \Lambda]\right\} \\ &= E[M_{N(t)|\Lambda}(s)] \\ &= E[M_{Poisson(\Lambda t)}(s)] \\ &\stackrel{form.}{=} E[e^{\Lambda t(e^s-1)}] \\ &= M_\Lambda[t(e^s-1)] \\ &= M_{Gamma(\alpha,\beta)}[t(e^s-1)] \\ &\stackrel{form.}{=} \left[\frac{\beta}{\beta - t(e^s-1)}\right]^\alpha, \ \beta > t(e^s-1). \end{aligned}$$

Group 2 — Renewal Processes **8.0 points**

1. *Suppose machines 1 and 2 process jobs independently. Moreover, admit processing times by machines 1 and 2 have Gamma($\alpha = 4, \lambda = 2$) and a Uniform$(0,4)$ distributions, respectively.*

   Obtain an approximate value to the probability that the two machines together process at least 90 jobs by time $t = 100$. (2.5)

- **Renewal processes**

  $\{N^{(1)}(t) : t \geq 0\} \perp\!\!\!\perp \{N^{(2)}(t) : t \geq 0\}$

  $N^{(j)}(t)$ = number of jobs processed by machine j until time t, $j = 1, 2$

- **Inter-renewal times**

$X_i^{(1)} \stackrel{i.i.d.}{\sim} X^{(1)},\ i \in \mathbb{N}$	$X_i^{(2)} \stackrel{i.i.d.}{\sim} X^{(2)},\ i \in \mathbb{N}$
$X^{(1)} \sim \text{Gamma}(\alpha = 4, \lambda = 2)$	$X^{(2)} \sim \text{Uniform}(a = 0, b = 4)$
$\mu^{(1)} = E[X^{(1)}] = \frac{4}{2} = 2$	$\mu^{(2)} = E[X^{(2)}] = \frac{0+4}{2} = 2$
$(\sigma^{(1)})^2 = V[X^{(1)}] = \frac{4}{2^2} = 1$	$(\sigma^{(2)})^2 = V[X^{(2)}] = \frac{(4-0)^2}{12} = \frac{4}{3}$

- **Approximate distributions**

  For large t,

$$N^{(j)}(t) \stackrel{a}{\sim}_{indep.} \text{Normal}\left(\frac{t}{\mu^{(j)}}, \frac{t(\sigma^{(j)})^2}{(\mu^{(j)})^3}\right), \ j = 1, 2.$$

  Consequently,[2]

$$N^{(1)}(t) + N^{(2)}(t) \stackrel{a}{\sim}_{indep.} \text{Normal}\left(\sum_{j=1}^2 \frac{t}{\mu^{(j)}}, \sum_{j=1}^2 \frac{t(\sigma^{(j)})^2}{(\mu^{(j)})^3}\right).$$

- **Requested probability (approximate value)**

$$\begin{aligned} P[N^{(1)}(t) + N^{(2)}(t) \geq n] &= 1 - P[N^{(1)}(t) + N^{(2)}(t) < n] \\ &\stackrel{form.}{\simeq} 1 - \Phi\left[\frac{n - \sum_{j=1}^2 \frac{t}{\mu^{(j)}}}{\sqrt{\sum_{j=1}^2 \frac{t(\sigma^{(j)})^2}{(\mu^{(j)})^3}}}\right] \\ &\stackrel{n=90,t=100,etc.}{=} 1 - \Phi\left[\frac{90 - (50+50)}{\sqrt{\frac{25}{2} + \frac{50}{3}}}\right] \\ &= 1 - \Phi\left(-\frac{10}{\sqrt{\frac{175}{6}}}\right) \\ &= \Phi\left(\frac{10}{\sqrt{\frac{175}{6}}}\right) \\ &\simeq \Phi(1.85) \\ &\stackrel{tables}{=} 0.9678. \end{aligned}$$

2. *Airplanes take off from an airport according to a renewal process with inter-renewal times with Gamma($\alpha = 2, \lambda = 1$) distribution.*

   (a) *What is the long-run rate at which take offs occur? Interpret it.* (1.0)

   - **Renewal process**

     $\{N(t) : t \geq 0\}$

     $N(t)$ = number of airplanes that took off until time t

   - **Inter-renewal times**

     $X_i \stackrel{i.i.d.}{\sim} X,\ i \in \mathbb{N}$

     $X \sim \text{Gamma}(\alpha = 2, \lambda = 1)$

     $\mu = E(X) = \frac{2}{1} = 2$

[2] Recall that the sum of independent normal distributions is normally distributed, etc.

- **Long-run rate at which airplanes take off**
  According to the SLLN for renewal processes (see formulae!),

$$\lim_{t\to+\infty} \frac{N(t)}{t} \stackrel{w.p.1}{=} \frac{1}{\mu} = \frac{1}{2}.$$

- **Interpretation**
  In the long-run one airplane takes off every two time units.

(b) *Derive the renewal function $m(t)$ of this renewal process, by using the Laplace-Stieltjes transform method and capitalizing on the table of important Laplace transforms in the formulae.* (2.5)

- **Deriving the renewal function**
  Since the inter-renewal times are continuous r.v. the LST of the inter-renewal distribution is given by

$$\begin{aligned}\tilde{F}(s) &= \int_{0^-}^{+\infty} e^{-sx}\, dF(x)\\ &= E(e^{-sX})\\ &= M_X(-s)\\ &\stackrel{form.,\,\alpha=2,\,\lambda=1}{=} \frac{1}{(1+s)^2}.\end{aligned}$$

Moreover, the LST of the renewal function can be obtained in terms of the one of F:

$$\begin{aligned}\tilde{m}(s) &\stackrel{form}{=} \frac{\tilde{F}(s)}{1-\tilde{F}(s)}\\ &= \frac{1}{(1+s)^2} \times \frac{1}{1-\frac{1}{(1+s)^2}}\\ &= \frac{1}{s(s+2)}.\end{aligned}$$

Taking advantage of the LT in the formulae, we successively get:

$$\begin{aligned}\frac{dm(t)}{dt} &= LT^{-1}\left[\tilde{m}(s), t\right]\\ &= LT^{-1}\left[\frac{1}{s(s+2)}, t\right]\\ &= \frac{e^{-0\times t}-e^{-2\times t}}{2-0}\\ &= \frac{1-e^{-2\times t}}{2}\\ m(t) &= \int_0^t \frac{1-e^{-2\times s}}{2}\, ds\\ &= \left.\left(\frac{s}{2}+\frac{e^{-2\times s}}{4}\right)\right|_0^t\\ &= \frac{t}{2}+\frac{e^{-2\times t}}{4}-\frac{1}{4}.\end{aligned}$$

(c) *Show that the renewal function obtained in (b) verifies the elementary renewal theorem.* (1.0)

- Verification of the elementary renewal theorem (ERT)

$$\begin{aligned}\lim_{t\to+\infty}\frac{m(t)}{t} &= \lim_{t\to+\infty}\frac{\frac{t}{2}+\frac{e^{-2\times t}}{4}-\frac{1}{4}}{t}\\ &= \frac{1}{2}\\ &= \frac{1}{\mu},\end{aligned}$$

  thus, verifying the ERT.

(d) *Admit Clotilde arrived to the airport at time $t = 100$. Compute the expected time until the first take off occurs after her arrival.* (1.0)

- **R.v.**
  $Y(t) \stackrel{form.}{=} S_{N(t)+1} - t =$ time until the first take off occurs after Clotilde's arrival at time t
- **Requested expected value**

$$\begin{aligned}E[Y(t)] &= E[S_{N(t)+1}]\\ &\stackrel{form.}{=} \mu[m(t)+1]-t\\ &\stackrel{(b),t=100}{=} 2\times\left[\left(\frac{100}{2}+\frac{e^{-2\times 100}}{4}-\frac{1}{4}\right)+1\right]-100\\ &\simeq 2\times 50.75-100\\ &= 1.5.\end{aligned}$$

- **Obs.**
  Since $t = 100$ is sufficiently large, $E[Y(100)] \stackrel{form.}{\simeq} \frac{E(X^2)}{2E(X)} = \frac{V(X)+E^2(X)}{2E(X)} = \cdots = 1.5$