Department of Mathematics, IST — Probability and Statistics Unit

Introduction to Stochastic Processes

1st. Test — 2nd. Semester — 2013/14
Duration: 1h30m — 2014/04/15 — 5PM, Room C9

- 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

Consider $a \neq 0$, $\omega \in \mathbb{R}$ and $\Theta \sim \text{Uniform}(-\pi, \pi)$, and admit $X(t) = a\cos(\omega t + \Theta)$ represents the cash flow into stock funds (measured as a percentage of total assets) at time t.

(a) Consider the stochastic process $\{X(t) : t \geq 0\}$ and obtain its mean function. (1.0)

- **Stochastic process**
  $\{X(t) : t \geq 0\}$
  $X(t) = a\cos(\omega t + \Theta) =$ cash flow... at time t
  $a \neq 0$, $\omega \in \mathbb{R}$, $\Theta \sim \text{Uniform}(-\pi, \pi)$
- **P.d.f. of Θ**
  $$f_\Theta(\theta) = \begin{cases} \frac{1}{2\pi}, & -\pi < \theta < \pi \\ 0, & \text{otherwise} \end{cases}$$
- **Mean function**
  $$\begin{aligned} E[X(t)] &= E[a\cos(\omega t + \Theta)] \\ &= a \times \int_{-\pi}^{\pi} \cos(\omega t + \theta) \times \frac{1}{2\pi}\, d\theta \\ &= \frac{a}{2\pi} \times \sin(\omega t + \theta)\big|_{-\pi}^{\pi} \\ &= \frac{a}{2\pi} \times [\sin(\omega t + \pi) - \sin(\omega t - \pi)] \\ &= \frac{a}{2\pi} \times [\sin(\omega t) - \sin(\omega t)] \\ &= 0 \end{aligned}$$

(b) Are we dealing with a second order weakly stationary process? (1.5)
Hint: Note that $\int_{-\pi}^{\pi} \cos(a + x)\cos(a + y + x)\, dx = \pi\cos(y)$.

- **Checking whether the process is (second order weakly) stationary**
  The autocovariance function of $\{X(t) : t \geq 0\}$ is, for $t, s \geq 0$, equal to
  $$\begin{aligned} cov\,(X(t), X(t+s)) &= E[X(t) \times X(t+s)] - E[X(t)] \times E[X(t+s)] \\ &\overset{(a)}{=} E[X(t) \times X(t+s)] \\ &= a^2 E\{\cos(\omega t + \Theta) \times \cos[\omega(t+s) + \Theta]\} \\ &= \frac{a^2}{2\pi} \times \int_{-\pi}^{\pi} \cos(\omega t + \theta) \times \cos[\omega t + \omega s + \theta]\, d\theta \end{aligned}$$

$$\begin{aligned} cov\,(X(t), X(t+s)) &\overset{a=\omega t,\, x=\theta,\, y=\omega s}{=} \frac{a^2}{2\pi} \times \pi\cos(\omega s) \\ &= \frac{a^2}{2}\cos(\omega s). \end{aligned}$$

Since $E[X(t)]$ does not depend on t and $cov\,(X(t), X(t+s))$ only depends on the time lag s, $\{X(t) : t \geq 0\}$ is a second order weakly stationary process.

Group 1 — Poisson Processes — 9.0 points

1. Suppose that the number of typographical errors — in a running text on a LED display — occur according with a Poisson process with rate $\lambda = 2$ (errors per hour).

   (a) Find the probability that exactly five typographical errors occur in the first three hours, given that in the first hour the running text had at least one typographical error. (1.5)

   - **Stochastic process**
     $\{N(t) : t \geq 0\} \sim PP(\lambda)$
     $N(t) =$ number of typographical errors by time t (time in hours)
   - **Relevant distributions**
     $N(t) \sim \text{Poisson}(\lambda t)$
     $(N(s) \mid N(t) = n) \sim \text{Binomial}(n, s/t)$, $0 < s < t$ (see formulae)
   - **Requested probability**
     $$\begin{aligned} P[N(3) = 5 \mid N(1) > 0] &\overset{Bayes'\ theo.}{=} \frac{P[N(1) > 0 \mid N(3) = 5] \times P[N(3) = 5]}{P[N(1) > 0]} \\ &= \frac{\{1 - P[N(1) = 0 \mid N(3) = 5]\} \times P[N(3) = 5]}{1 - P[N(1) = 0]} \\ &= \frac{\left[1 - \binom{5}{0}(1/3)^0(1 - 1/3)^{5-0}\right] \times \frac{e^{-6}6^5}{5!}}{1 - \frac{e^{-2}2^0}{0!}} \\ &= \frac{[1 - (2/3)^5] \times \frac{e^{-6}6^5}{5!}}{1 - e^{-2}} \\ &\simeq 0.161301. \end{aligned}$$

     [Alternatively,
     $$\begin{aligned} P[N(3) = 5 \mid N(1) > 0] &= \frac{P[N(1) > 0,\, N(3) = 5]}{P[N(1) > 0]} \\ &= \frac{P[N(3) = 5] - P[N(1) = 0,\, N(3) = 5]}{1 - P[N(1) = 0]} \\ &\overset{indep.\ incr.}{=} \frac{P[N(3) = 5] - P[N(1) = 0] \times P[N(3) - N(1) = 5]}{1 - P[N(1) = 0]} \\ &\overset{station.\ incr.}{=} \frac{P[N(3) = 5] - P[N(1) = 0] \times P[N(3-1) = 5]}{1 - P[N(1) = 0]} \\ &\overset{N(t) \sim Poi(2t)}{=} \frac{\frac{e^{-2\times 3}(2\times 3)^5}{5!} - \frac{e^{-2}2^0}{0!} \times \frac{e^{-2\times 2}(2\times 2)^5}{5!}}{1 - \frac{e^{-2}2^0}{0!}} \\ &= \frac{\frac{e^{-6}}{5!} \times (6^5 - 4^5)}{1 - e^{-2}} \\ &\simeq 0.161301.] \end{aligned}$$

(b) Two proofreaders independently read the running text. Suppose that each error is independently found by proofreader i with probability p_i, $i = 1, 2$. (1.5)

Let $\{N_i(t) : t \geq 0\}$, $i = 1, 2, 3, 4$, denote the number of errors that are found until time t by

(1) proofreader 1 but not by proofreader 2,

(2) proofreader 2 but not by proofreader 1,

(3) both proofreaders,

(4) neither proofreader,

respectively.

What are the marginal distributions and the joint p.f. of these r.v.?

- **Split processes**
  The original PP with rate λ, $\{N(t) : t \geq 0\}$, is now split in four other processes, $\{N_i(t) : t \geq 0\}$ $(i = 1, 2, 3, 4)$, referring to the counts of four distinct types of events.
- **Requested marginal distributions and joint p.f.**
  The resulting processes are INDEPENDENT and also POISSON. Moreover, since the two proofreaders independently read the running text, the associated rates are

$$\begin{aligned} \lambda_1 &= \lambda\, p_1(1-p_2), \\ \lambda_2 &= \lambda\, p_2(1-p_1), \\ \lambda_3 &= \lambda\, p_1 p_2, \\ \lambda_4 &= \lambda\,(1-p_1)(1-p_2), \end{aligned}$$

respectively. Consequently,

$$N_i(t) \sim_{indep.} \text{Poisson}(\lambda_i\, t),\ i = 1, 2, 3, 4.$$

Furthermore, the joint p.f. of these independent r.v. is the product of their marginal p.f.:

$$P[N_i(t) = n_i,\ i = 1, 2, 3, 4] \overset{N_i(t) \sim_{indep.} Poisson(\lambda_i\, t)}{=} \prod_{i=1}^{4} \frac{e^{-\lambda_i\, t} \times (\lambda_i\, t)^{n_i}}{n_i!},\ n_i \in \mathbb{N}_0.$$

2. Policyholders of a large insurance company have accidents according to a Poisson process with rate $\lambda = 5$ (accidents per day). The time R it takes a policyholder to report an accident is exponentially distributed with mean equal to $\mu = 10$ days. (2.0)

Find the probability there are exactly $n = 10$ unreported accidents by time $t = 2$ (days).

- **Stochastic process**
  $\{N(t) : t \geq 0\} \sim PP(\lambda = 5)$
  $N(t) =$ number of accidents by day t
- **Non-homogenous Bernoulli splitting**
  An accident, which occurred at time s $(0 < s < t)$, remains unreported at time t with probability

$$\begin{aligned} p(s) &= P(R > t - s) \\ &= \overline{F}_{Exp(\mu^{-1})}(t - s) \\ &= e^{-\mu^{-1}(t-s)}. \end{aligned}$$

Then the number of unreported accidents by day t, $N_U(t)$, results from a non-homogenous Bernoulli splitting of $\{N(t) : t \geq 0\}$. As a consequence,

$$N_U(t) \overset{form.}{\sim} \text{Poisson}\left(\lambda \int_0^t p(s)\, ds = \lambda \int_0^t e^{-\mu^{-1}(t-s)}\, ds = \lambda\, \mu\, (1 - e^{-\mu^{-1} t}).\right)$$

- **Requested probability**

$$\begin{aligned} P[N_U(2) = 10] &= \frac{e^{-5\times 10\times(1-e^{-10^{-1}\times 2})}\, [5 \times 10 \times (1 - e^{-10^{-1}\times 2})]^{10}}{10!} \\ &\simeq 0.119390. \end{aligned}$$

[Alternatively, since $5 \times 10\,(1 - e^{-10^{-1}\times 2}) \simeq 9.063462 \simeq 9$, we could have used the tables and add that $P[N_U(2) = 10] = F_{Poi(9)}(10) - F_{Poi(9)}(9) \overset{tables}{\simeq} 0.7060 - 0.5874 = 0.1186$.]

3. Suppose that workmen incur accidents in accordance with a non-homogeneous Poisson process with mean value function $m(t) = t^2 + 2t$, $t \geq 0$. (1.5)

Compute the probability that the 5^{th} accident occurs in the interval $(2, 4]$.

- **Stochastic process**
  $\{N(t) : t \geq 0\} \sim NHPP$
  $N(t) =$ number of accidents by day t
- **Mean value function**
  $m(t) = t^2 + 2t$, $t \geq 0$
- **R.v.**
  $S_5 =$ time of the 5^{th} arrival
- **Requested probability**

$$\begin{aligned} P(2 < S_5 \leq 4) &= P(S_5 \leq 4) - P(S_5 \leq 2) \\ &\overset{form.}{=} P[N(4) \geq 5] - P[N(2) \geq 5] \\ &= \{1 - P[N(4) \leq 4]\} - \{1 - P[N(2) \leq 4]\} \\ &= P[N(2) \leq 4] - P[N(4) \leq 4] \\ &\overset{N(t)\sim Poisson(m(t))}{=} F_{Poisson(m(2))}(4) - F_{Poisson(m(4))}(4) \\ &= F_{Poisson(2^2+2\times 2)}(4) - F_{Poisson(4^2+2\times 4)}(4) \\ &= F_{Poisson(8)}(4) - F_{Poisson(24)}(4) \\ &\overset{tables}{=} 0.0996 - 0.0000 \\ &= 0.0996. \end{aligned}$$

4. In good years, storms occur according to a Poisson process with rate 3 per time unit, while in bad years they occur according to a Poisson process with rate 5 per time unit. Moreover, suppose next year will be good (resp. bad) with probability 0.3 (resp. 0.7).

Let $N(t)$ denote the number of storms during the first t time units of next year.

(a) What sort of stochastic process is $\{N(t) : t \geq 0\}$? Derive $P[N(t) = n]$. (1.5)

- **Stochastic process**
  $\{N(t) : t \geq 0\} \sim ConditionalPP(\Lambda)$
  $N(t)$ = number of storms during the first t time units of next year
- **Random arrival rate**
  Λ is a r.v. with p.f.

$$P(\Lambda = \lambda) = \begin{cases} 0.3, & \lambda = 3 \text{ (good year)} \\ 0.7, & \lambda = 5 \text{ (bad year)} \\ 0, & \text{otherwise.} \end{cases}$$

- **P.f. of N(t) conditional to $\Lambda = \lambda$**
  Since for a conditional Poisson process, $(N(t) \mid \Lambda = \lambda) \sim \text{Poisson}(\lambda t)$, we have

$$P[N(t) = n \mid \Lambda = \lambda] = \frac{e^{-\lambda t}(\lambda t)^n}{n!}, \; n \in \mathbb{N}_0.$$

- **Requested p.f.**
  By the total probability law,

$$\begin{aligned} P[N(t) = n] &= P[N(t) \mid \Lambda = 3] \times P(\Lambda = 3) + P[N(t) \mid \Lambda = 5] \times P(\Lambda = 5) \\ &= \frac{e^{-3t}(3t)^n}{n!} \times 0.3 + \frac{e^{-5t}(5t)^n}{n!} \times 0.7, \; n \in \mathbb{N}_0. \end{aligned}$$

  [Alternatively, we could have used the formulae and written:

$$\begin{aligned} P[N(t) = n] &= \int_0^{+\infty} \frac{e^{-\lambda t}(\lambda t)^n}{n!} \, dG_\Lambda(\lambda) \\ &= \sum_\lambda \frac{e^{-\lambda t}(\lambda t)^n}{n!} \times P(\Lambda = \lambda) \\ &= \frac{e^{-3t}(3t)^n}{n!} \times 0.3 + \frac{e^{-5t}(5t)^n}{n!} \times 0.7, \; n \in \mathbb{N}_0.] \end{aligned}$$

(b) If next year starts off with two storms by time $t = 1$, what is the conditional probability it is a good year? (1.0)

- **Requested probability**
  Applying Bayes' theorem leads to

$$\begin{aligned} P[\Lambda = 3 \mid N(1) = 2] &= \frac{P[N(1) = 2 \mid \Lambda = 3] \times P(\Lambda = 3)}{P[N(1) = 2]} \\ &= \frac{\frac{e^{-3}3^2}{2!} \times 0.3}{\frac{e^{-3}3^2}{2!} \times 0.3 + \frac{e^{-5}5^2}{2!} \times 0.7} \\ &\simeq 0.532716. \end{aligned}$$

Group 2 — Renewal Processes **8.5 points**

1. Consider a train station to which passengers arrive in accordance with a renewal process and admit that the inter-renewal times (in seconds) have Weibull$(\alpha = 1, \beta = \frac{1}{3})$ distribution.

(a) Calculate an approximate value to the probability that at least 600 passengers arrive in the first hour. (1.5)

- **Renewal process**
  $\{N(t) : t \geq 0\}$
  $N(t)$ = number of passengers that arrived by time t
- **Inter-renewal times**
  $X_i \stackrel{i.i.d.}{\sim} X, \; i \in \mathbb{N}$
  $X \sim \text{Weibull}(\alpha = 1, \beta = \frac{1}{3})$
  $\mu = E(X) \stackrel{form.}{=} \alpha \times \Gamma(1 + 1/\beta) = \Gamma(4) = (4-1)! = 6$
  $\sigma^2 = V(X) \stackrel{form.}{=} \alpha^2 \times [\Gamma(1 + 2/\beta) - \Gamma^2(1 + 1/\beta)] = \Gamma(7) - 6^2 = 6! - 36 = 684$
- **Requested approximate probability**

$$\begin{aligned} P[N(t) \geq n] &= 1 - P[N(t) < n] \\ &\stackrel{form.}{\simeq} 1 - \Phi\left(\frac{n - t/\mu}{\sqrt{t\sigma^2/\mu^3}}\right) \\ &\stackrel{t=3600,\, n=600}{=} 1 - \Phi\left(\frac{600 - 3600/6}{\sqrt{3600 \times 684/6^3}}\right) \\ &= 1 - \Phi(0) \\ &= 0.5. \end{aligned}$$

(b) Admit an officer inspected the train station at 5PM.

(i) Provide an approximation to the expected value of the time until the first arrival after this inspection. (1.5)

- **Recurrence time**
  $Y(t) \stackrel{form.}{=} S_{N(t)+1} - t$ = time until the first arrival after the inspection at time t
- **Requested approximate expected value**
  Since the value of $t = (12 + 5) \times 3600 = 61200$ sec. is rather large,

$$\begin{aligned} E[Y(t)] &\simeq \lim_{z \to +\infty} E[Y(z)] \\ &\stackrel{form.}{=} \frac{E(X^2)}{2E(X)} \\ &\stackrel{(a)}{=} \frac{684 + 6^2}{2 \times 6} \\ &= 60 \, \text{sec.} \end{aligned}$$

(ii) Provide an approximate value to the probability that the last arrival before this inspection occurred more than 27 seconds ago. (2.5)

- **C.d.f. of the inter-renewal times**
  $X \sim \text{Weibull}(\alpha = 1, \beta = \frac{1}{3})$
  $f(x) \stackrel{form.}{=} \frac{1}{3} x^{\frac{1}{3}-1} e^{-x^{\frac{1}{3}}}, \; x \geq 0$
  $F(x) = \int_0^x f(y)\, dy = 1 - e^{-x^{\frac{1}{3}}}, \; x \geq 0$
- **Another recurrence time**
  $A(t) \stackrel{form}{=} t - S_{N(t)}$ = time until the last arrival before the inspection at time t
- **Requested probability** (approximate value)
  We can once again invoke that $t = (12 + 5) \times 3600 = 61200$ sec. is sufficiently large and provide the following approximate value

$$
\begin{aligned}
P[A(t) > x] &\simeq 1 - \lim_{z\to+\infty} P[A(z) \le x] \\
&\overset{form.}{=} 1 - \frac{\int_0^x [1 - F(u)]\,du}{E(X)} \\
&= 1 - \frac{\int_0^x e^{-u^{\frac{1}{3}}}\,du}{E(X)}.
\end{aligned}
$$

If we change variables ($w = u^{\frac{1}{3}};\ u = w^3;\ du = 3w^2\,dw$) then we get

$$
\begin{aligned}
\int_0^x e^{-u^{\frac{1}{3}}}\,du &= \int_0^{x^{\frac{1}{3}}} 3\,w^{3-1} e^{-w}\,dw \\
&= 3 \times \Gamma(3) \int_0^{x^{\frac{1}{3}}} f_{Gamma(3,1)}(w)\,dw \\
&= 6 \times F_{Gamma(n=3,\lambda=1)}(x^{1/3}) \\
&\overset{form.}{=} 6 \times \left[1 - F_{Poisson(1\times x^{1/3})}(3-1)\right]
\end{aligned}
$$

and

$$
\begin{aligned}
P[A(61200) > 27] &\simeq 1 - \frac{6 \times \left[1 - F_{Poisson(27^{1/3})}(2)\right]}{6} \\
&= F_{Poisson(3)}(2) \\
&\overset{tables}{=} 0.4232.
\end{aligned}
$$

2. Consider a renewal reward process $\{R(t) = \sum_{n=1}^{N(t)} R_n : t \ge 0\}$.

   (a) By conditioning on the epoch of the last renewal prior to t, show that $E\left[R_{N(t)+1}\right]$ satisfies the following renewal-type equation (1.5)

   $$E\left[R_{N(t)+1}\right] = h(t) + \int_0^t h(t-s)\,dm(s),$$

   where $h(y) = \int_y^{+\infty} E(R \mid X = x)\,dF(x)$.

   Note: $S_{N(t)}$ is a mixed r.v. such that $P[S_{N(t)} = 0] = P(X > t) = \bar{F}(t)$ and $dF_{S_{N(t)}} = \bar{F}(t-s)\,dm(s)$.

   - **Reward renewal process**
     $\{R(t) = \sum_{n=1}^{N(t)} R_n : t \ge 0\}$
     $\{N(t) : t \ge 0\}$ is a renewal process with IRT $X_n \overset{i.i.d.}{\sim} F$, $n \in \mathbb{N}$ (F not lattice)
     $R_n =$ reward at the n^{th} renewal
     $(X_n, R_n) \overset{i.i.d.}{\sim} (X, R)$, $n \in \mathbb{N}$
   - **Deriving** $E\left[R_{N(t)+1}\right]$
     Recall that the epoch of the last renewal prior to t, $S_{N(t)}$, is a mixed r.v. such that $P[S_{N(t)} = 0] = P(X > t) = \bar{F}(t)$ and $dF_{S_{N(t)}}(s) = \bar{F}(t-s)\,dm(s)$.
     Conditioning on $S_{N(t)} = s$ $(0 \le s < t)$, we obtain

$$
\begin{aligned}
E\left[R_{N(t)+1}\right] &= E\left\{E\left[R_{N(t)+1} \mid S_{N(t)}\right]\right\} \\
&= E\left[R_{N(t)+1} \mid S_{N(t)} = 0\right] \times P[S_{N(t)} = 0] \\
&\quad + \int_0^t E\left[R_{N(t)+1} \mid S_{N(t)} = s\right]\,dF_{S_{N(t)}}(s)
\end{aligned}
$$

$$
\begin{aligned}
E\left[R_{N(t)+1}\right] &= E(R_1 \mid X_1 > t) \times P(X_1 > t) \\
&\quad + \int_0^t E(R_1 \mid X_1 > t-s)\,\bar{F}(t-s)\,dm(s) \\
&= \int_t^{+\infty} E(R \mid X = x)\,dF(x) \\
&\quad + \int_0^t \left\{\int_{t-s}^{+\infty} E(R \mid X = x)\,dF(x)\right\} dm(s) \\
&= h(t) + \int_0^t h(t-s)\,dm(s),
\end{aligned}
$$

where $h(y) = \int_y^{+\infty} E(R \mid X = x)\,dF(x)$. QED

- **[Obs.**

$$
\begin{aligned}
E(R \mid X > t) \times P(X > t) &= \left[\int_{-\infty}^{+\infty} r\,dF_{R|X>t}(r)\right] \times P(X > t) \\
&= \left[\int_{-\infty}^{+\infty} r\,\frac{dP(R \le r, X > t)}{P(X > t)}\right] \times P(X > t) \\
&= \int_{-\infty}^{+\infty} r\,d\left[\int_t^{+\infty} \int_{-\infty}^{r} dF_{R,X}(u,x)\right] \\
&= \int_t^{+\infty} \int_{-\infty}^{+\infty} r\,dF_{R,X}(r,x) \\
&= \int_t^{+\infty} \int_{-\infty}^{+\infty} r\,\frac{dF_{R,X}(r,x)}{dF(x)}\,dF(x) \\
&= \int_t^{+\infty} \left[\int_{-\infty}^{+\infty} r\,dF_{R|X=x}(r)\right] dF(x) \\
&= \int_t^{+\infty} E(R \mid X = x)\,dF(x).]
\end{aligned}
$$

   (b) Prove that (1.5)

   $$\lim_{t\to+\infty} E\left[R_{N(t)+1}\right] = \frac{E(R \times X)}{E(X)}.$$

   Note: In this proof assume the inter-renewal distribution F is not lattice and that any relevant function is dRi.

   - **Proof**

$$
\begin{aligned}
\lim_{t\to+\infty} E\left[R_{N(t)+1}\right] &\overset{KRT}{=} \frac{\int_0^{+\infty} h(y)\,dy}{E(X)} \\
&= \frac{1}{E(X)} \times \int_0^{+\infty} \left[\int_y^{+\infty} E(R \mid X = x)\,dF_X(x)\right] dy \\
&= \frac{1}{E(X)} \times \int_0^{+\infty} \left[E(R \mid X = x) \times \int_0^x dy\right] dF_X(x) \\
&= \frac{1}{E(X)} \times \int_0^{+\infty} \left\{\left[\int_{-\infty}^{+\infty} r\,dF_{R|X=x}(r)\right] \times x\right\} dF_X(x) \\
&= \frac{1}{E(X)} \times \int_0^{+\infty} \int_{-\infty}^{+\infty} r\,x\,dF_{R,X}(r,x) \\
&= \frac{E(R \times X)}{E(X)}.
\end{aligned}
$$