20.2 Poisson Processes

Poisson process

- $\{X_t : t \in [0, \infty)\}$ = number of events up to and including time t
- $X_0 = 0$
- Independent increments:

$$\forall t_0 < \cdots < t_n : X_{t_1} - X_{t_0} \perp\!\!\!\perp \cdots \perp\!\!\!\perp X_{t_n} - X_{t_{n-1}}$$

- Intensity function $\lambda(t)$
  - $\mathbb{P}[X_{t+h} - X_t = 1] = \lambda(t)h + o(h)$
  - $\mathbb{P}[X_{t+h} - X_t = 2] = o(h)$
- $X_{s+t} - X_s \sim \text{Po}(m(s+t) - m(s))$ where $m(t) = \int_0^t \lambda(s)\,ds$

Homogeneous Poisson process

$$\lambda(t) \equiv \lambda \implies X_t \sim \text{Po}(\lambda t) \qquad \lambda > 0$$

Waiting times

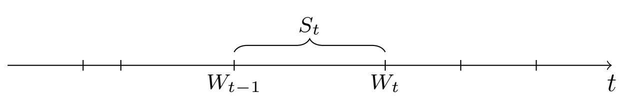
$$W_t := \text{time at which } X_t \text{ occurs}$$

$$W_t \sim \text{Gamma}\left(t, \frac{1}{\lambda}\right)$$

Interarrival times

$$S_t = W_{t+1} - W_t$$

$$S_t \sim \text{Exp}\left(\frac{1}{\lambda}\right)$$



21 Time Series

Mean function

$$\mu_{x_t} = \mathbb{E}[x_t] = \int_{-\infty}^{\infty} x f_t(x)\,dx$$

Autocovariance function

$$\gamma_x(s,t) = \mathbb{E}[(x_s - \mu_s)(x_t - \mu_t)] = \mathbb{E}[x_s x_t] - \mu_s \mu_t$$

$$\gamma_x(t,t) = \mathbb{E}\left[(x_t - \mu_t)^2\right] = \mathbb{V}[x_t]$$

Autocorrelation function (ACF)

$$\rho(s,t) = \frac{\text{Cov}[x_s, x_t]}{\sqrt{\mathbb{V}[x_s]\,\mathbb{V}[x_t]}} = \frac{\gamma(s,t)}{\sqrt{\gamma(s,s)\gamma(t,t)}}$$

Cross-covariance function (CCV)

$$\gamma_{xy}(s,t) = \mathbb{E}[(x_s - \mu_{x_s})(y_t - \mu_{y_t})]$$

Cross-correlation function (CCF)

$$\rho_{xy}(s,t) = \frac{\gamma_{xy}(s,t)}{\sqrt{\gamma_x(s,s)\gamma_y(t,t)}}$$

Backshift operator

$$B^k(x_t) = x_{t-k}$$

Difference operator

$$\nabla^d = (1 - B)^d$$

White noise

- $w_t \sim wn(0, \sigma_w^2)$
- Gaussian: $w_t \overset{iid}{\sim} \mathcal{N}(0, \sigma_w^2)$
- $\mathbb{E}[w_t] = 0 \quad t \in T$
- $\mathbb{V}[w_t] = \sigma^2 \quad t \in T$
- $\gamma_w(s,t) = 0 \quad s \neq t \ \wedge \ s,t \in T$

Random walk

- Drift δ
- $x_t = \delta t + \sum_{j=1}^t w_j$
- $\mathbb{E}[x_t] = \delta t$

Symmetric moving average

$$m_t = \sum_{j=-k}^{k} a_j x_{t-j} \qquad \text{where } a_j = a_{-j} \geq 0 \text{ and } \sum_{j=-k}^{k} a_j = 1$$

21.1 Stationary Time Series

Strictly stationary

$$\mathbb{P}\left[x_{t_1} \le c_1, \ldots, x_{t_k} \le c_k\right] = \mathbb{P}\left[x_{t_1+h} \le c_1, \ldots, x_{t_k+h} \le c_k\right]$$

$$\forall k \in \mathbb{N}, t_k, c_k, h \in \mathbb{Z}$$

Weakly stationary

- $\mathbb{E}\left[x_t^2\right] < \infty \qquad \forall t \in \mathbb{Z}$
- $\mathbb{E}\left[x_t^2\right] = m \qquad \forall t \in \mathbb{Z}$
- $\gamma_x(s,t) = \gamma_x(s+r, t+r) \qquad \forall r,s,t \in \mathbb{Z}$

Autocovariance function

- $\gamma(h) = \mathbb{E}\left[(x_{t+h} - \mu)(x_t - \mu)\right] \qquad \forall h \in \mathbb{Z}$
- $\gamma(0) = \mathbb{E}\left[(x_t - \mu)^2\right]$
- $\gamma(0) \ge 0$
- $\gamma(0) \ge |\gamma(h)|$
- $\gamma(h) = \gamma(-h)$

Autocorrelation function (ACF)

$$\rho_x(h) = \frac{\text{Cov}\left[x_{t+h}, x_t\right]}{\sqrt{\mathbb{V}\left[x_{t+h}\right]\mathbb{V}\left[x_t\right]}} = \frac{\gamma(t+h,t)}{\sqrt{\gamma(t+h,t+h)\gamma(t,t)}} = \frac{\gamma(h)}{\gamma(0)}$$

Jointly stationary time series

$$\gamma_{xy}(h) = \mathbb{E}\left[(x_{t+h} - \mu_x)(y_t - \mu_y)\right]$$

$$\rho_{xy}(h) = \frac{\gamma_{xy}(h)}{\sqrt{\gamma_x(0)\gamma_y(h)}}$$

Linear process

$$x_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j w_{t-j} \quad \text{where} \quad \sum_{j=-\infty}^{\infty} |\psi_j| < \infty$$

$$\gamma(h) = \sigma_w^2 \sum_{j=-\infty}^{\infty} \psi_{j+h}\psi_j$$

21.2 Estimation of Correlation

Sample mean

$$\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t$$

Sample variance

$$\mathbb{V}\left[\bar{x}\right] = \frac{1}{n}\sum_{h=-n}^{n}\left(1 - \frac{|h|}{n}\right)\gamma_x(h)$$

Sample autocovariance function

$$\widehat{\gamma}(h) = \frac{1}{n}\sum_{t=1}^{n-h}(x_{t+h} - \bar{x})(x_t - \bar{x})$$

Sample autocorrelation function

$$\widehat{\rho}(h) = \frac{\widehat{\gamma}(h)}{\widehat{\gamma}(0)}$$

Sample cross-variance function

$$\widehat{\gamma}_{xy}(h) = \frac{1}{n}\sum_{t=1}^{n-h}(x_{t+h} - \bar{x})(y_t - \bar{y})$$

Sample cross-correlation function

$$\widehat{\rho}_{xy}(h) = \frac{\widehat{\gamma}_{xy}(h)}{\sqrt{\widehat{\gamma}_x(0)\widehat{\gamma}_y(0)}}$$

Properties

- $\sigma_{\widehat{\rho}_x(h)} = \frac{1}{\sqrt{n}}$ if x_t is white noise
- $\sigma_{\widehat{\rho}_{xy}(h)} = \frac{1}{\sqrt{n}}$ if x_t or y_t is white noise

21.3 Non-Stationary Time Series

Classical decomposition model

$$x_t = \mu_t + s_t + w_t$$

- μ_t = trend
- s_t = seasonal component
- w_t = random noise term

21.3.1 Detrending

Least squares

1. Choose trend model, e.g., $\mu_t = \beta_0 + \beta_1 t + \beta_2 t^2$
2. Minimize RSS to obtain trend estimate $\widehat{\mu}_t = \widehat{\beta}_0 + \widehat{\beta}_1 t + \widehat{\beta}_2 t^2$
3. Residuals $\triangleq$ noise w_t

Moving average

- The *low-pass* filter v_t is a symmetric moving average m_t with $a_j = \frac{1}{2k+1}$:

$$v_t = \frac{1}{2k+1} \sum_{i=-k}^{k} x_{t-1}$$

- If $\frac{1}{2k+1} \sum_{i=-k}^{k} w_{t-j} \approx 0$, a linear trend function $\mu_t = \beta_0 + \beta_1 t$ passes without distortion

Differencing

- $\mu_t = \beta_0 + \beta_1 t \implies \nabla x_t = \beta_1$

21.4 ARIMA models

Autoregressive polynomial

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z_p \qquad z \in \mathbb{C} \land \phi_p \neq 0$$

Autoregressive operator

$$\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$$

Autoregressive model order p, $\mathsf{AR}(p)$

$$x_t = \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + w_t \iff \phi(B) x_t = w_t$$

$\mathsf{AR}(1)$

- $x_t = \phi^k (x_{t-k}) + \sum_{j=0}^{k-1} \phi^j (w_{t-j}) \overset{k\to\infty, |\phi|<1}{=} \underbrace{\sum_{j=0}^{\infty} \phi^j (w_{t-j})}_{\text{linear process}}$
- $\mathbb{E}[x_t] = \sum_{j=0}^{\infty} \phi^j (\mathbb{E}[w_{t-j}]) = 0$
- $\gamma(h) = \mathrm{Cov}[x_{t+h}, x_t] = \frac{\sigma_w^2 \phi^h}{1-\phi^2}$
- $\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \phi^h$
- $\rho(h) = \phi \rho(h-1) \quad h = 1, 2, \ldots$

Moving average polynomial

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z_q \qquad z \in \mathbb{C} \land \theta_q \neq 0$$

Moving average operator

$$\theta(B) = 1 + \theta_1 B + \cdots + \theta_p B^p$$

$\mathsf{MA}(q)$ (moving average model order q)

$$x_t = w_t + \theta_1 w_{t-1} + \cdots + \theta_q w_{t-q} \iff x_t = \theta(B) w_t$$

$$\mathbb{E}[x_t] = \sum_{j=0}^{q} \theta_j \mathbb{E}[w_{t-j}] = 0$$

$$\gamma(h) = \mathrm{Cov}[x_{t+h}, x_t] = \begin{cases} \sigma_w^2 \sum_{j=0}^{q-h} \theta_j \theta_{j+h} & 0 \le h \le q \\ 0 & h > q \end{cases}$$

$\mathsf{MA}(1)$

$$x_t = w_t + \theta w_{t-1}$$

$$\gamma(h) = \begin{cases} (1+\theta^2)\sigma_w^2 & h = 0 \\ \theta \sigma_w^2 & h = 1 \\ 0 & h > 1 \end{cases}$$

$$\rho(h) = \begin{cases} \frac{\theta}{(1+\theta^2)} & h = 1 \\ 0 & h > 1 \end{cases}$$

$\mathsf{ARMA}(p, q)$

$$x_t = \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \cdots + \theta_q w_{t-q}$$

$$\phi(B) x_t = \theta(B) w_t$$

Partial autocorrelation function (PACF)

- $x_i^{h-1} \triangleq$ regression of x_i on $\{x_{h-1}, x_{h-2}, \ldots, x_1\}$
- $\phi_{hh} = corr(x_h - x_h^{h-1}, x_0 - x_0^{h-1}) \quad h \ge 2$
- E.g., $\phi_{11} = corr(x_1, x_0) = \rho(1)$

$\mathsf{ARIMA}(p, d, q)$

$$\nabla^d x_t = (1 - B)^d x_t \text{ is } \mathsf{ARMA}(p, q)$$

$$\phi(B)(1 - B)^d x_t = \theta(B) w_t$$

Exponentially Weighted Moving Average (EWMA)

$$x_t = x_{t-1} + w_t - \lambda w_{t-1}$$

$$x_t = \sum_{j=1}^{\infty}(1-\lambda)\lambda^{j-1}x_{t-j} + w_t \quad \text{when } |\lambda| < 1$$

$$\tilde{x}_{n+1} = (1-\lambda)x_n + \lambda\tilde{x}_n$$

Seasonal ARIMA

- Denoted by $\mathsf{ARIMA}\,(p,d,q) \times (P,D,Q)_s$
- $\Phi_P(B^s)\phi(B)\nabla_s^D\nabla^d x_t = \delta + \Theta_Q(B^s)\theta(B)w_t$

21.4.1 Causality and Invertibility

$\mathsf{ARMA}\,(p,q)$ is causal (future-independent) $\iff \exists\{\psi_j\} : \sum_{j=0}^{\infty}\psi_j < \infty$ such that

$$x_t = \sum_{j=0}^{\infty} w_{t-j} = \psi(B)w_t$$

$\mathsf{ARMA}\,(p,q)$ is invertible $\iff \exists\{\pi_j\} : \sum_{j=0}^{\infty}\pi_j < \infty$ such that

$$\pi(B)x_t = \sum_{j=0}^{\infty} X_{t-j} = w_t$$

Properties

- $\mathsf{ARMA}\,(p,q)$ causal $\iff$ roots of $\phi(z)$ lie outside the unit circle

$$\psi(z) = \sum_{j=0}^{\infty}\psi_j z^j = \frac{\theta(z)}{\phi(z)} \quad |z| \le 1$$

- $\mathsf{ARMA}\,(p,q)$ invertible $\iff$ roots of $\theta(z)$ lie outside the unit circle

$$\pi(z) = \sum_{j=0}^{\infty}\pi_j z^j = \frac{\phi(z)}{\theta(z)} \quad |z| \le 1$$

Behavior of the ACF and PACF for causal and invertible ARMA models

	$\mathsf{AR}\,(p)$	$\mathsf{MA}\,(q)$	$\mathsf{ARMA}\,(p,q)$
ACF	tails off	cuts off after lag q	tails off
PACF	cuts off after lag p	tails off q	tails off

21.5 Spectral Analysis

Periodic process

$$\begin{aligned} x_t &= A\cos(2\pi\omega t + \phi) \\ &= U_1\cos(2\pi\omega t) + U_2\sin(2\pi\omega t) \end{aligned}$$

- Frequency index ω (cycles per unit time), period $1/\omega$
- Amplitude A
- Phase ϕ
- $U_1 = A\cos\phi$ and $U_2 = A\sin\phi$ often normally distributed RV's

Periodic mixture

$$x_t = \sum_{k=1}^{q}\left(U_{k1}\cos(2\pi\omega_k t) + U_{k2}\sin(2\pi\omega_k t)\right)$$

- U_{k1}, U_{k2}, for $k = 1,\dots,q$, are independent zero-mean RV's with variances σ_k^2
- $\gamma(h) = \sum_{k=1}^{q}\sigma_k^2\cos(2\pi\omega_k h)$
- $\gamma(0) = \mathbb{E}\left[x_t^2\right] = \sum_{k=1}^{q}\sigma_k^2$

Spectral representation of a periodic process

$$\begin{aligned} \gamma(h) &= \sigma^2\cos(2\pi\omega_0 h) \\ &= \frac{\sigma^2}{2}e^{-2\pi i\omega_0 h} + \frac{\sigma^2}{2}e^{2\pi i\omega_0 h} \\ &= \int_{-1/2}^{1/2} e^{2\pi i\omega h}\,dF(\omega) \end{aligned}$$

Spectral distribution function

$$F(\omega) = \begin{cases} 0 & \omega < -\omega_0 \\ \sigma^2/2 & -\omega \le \omega < \omega_0 \\ \sigma^2 & \omega \ge \omega_0 \end{cases}$$

- $F(-\infty) = F(-1/2) = 0$
- $F(\infty) = F(1/2) = \gamma(0)$

Spectral density

$$f(\omega) = \sum_{h=-\infty}^{\infty}\gamma(h)e^{-2\pi i\omega h} \quad -\frac{1}{2} \le \omega \le \frac{1}{2}$$

- Needs $\sum_{h=-\infty}^{\infty}|\gamma(h)| < \infty \implies \gamma(h) = \int_{-1/2}^{1/2} e^{2\pi i\omega h}f(\omega)\,d\omega \quad h = 0, \pm 1, \dots$
- $f(\omega) \ge 0$
- $f(\omega) = f(-\omega)$
- $f(\omega) = f(1-\omega)$
- $\gamma(0) = \mathbb{V}\left[x_t\right] = \int_{-1/2}^{1/2} f(\omega)\,d\omega$
- White noise: $f_w(\omega) = \sigma_w^2$

- ARMA (p,q), $\phi(B)x_t = \theta(B)w_t$:

$$f_x(\omega) = \sigma_w^2 \frac{|\theta(e^{-2\pi i\omega})|^2}{|\phi(e^{-2\pi i\omega})|^2}$$

where $\phi(z) = 1 - \sum_{k=1}^{p} \phi_k z^k$ and $\theta(z) = 1 + \sum_{k=1}^{q} \theta_k z^k$

Discrete Fourier Transform (DFT)

$$d(\omega_j) = n^{-1/2} \sum_{i=1}^{n} x_t e^{-2\pi i \omega_j t}$$

Fourier/Fundamental frequencies

$$\omega_j = j/n$$

Inverse DFT

$$x_t = n^{-1/2} \sum_{j=0}^{n-1} d(\omega_j) e^{2\pi i \omega_j t}$$

Periodogram

$$I(j/n) = |d(j/n)|^2$$

Scaled Periodogram

$$\begin{aligned} P(j/n) &= \frac{4}{n} I(j/n) \\ &= \left(\frac{2}{n}\sum_{t=1}^{n} x_t \cos(2\pi t j/n)\right)^2 + \left(\frac{2}{n}\sum_{t=1}^{n} x_t \sin(2\pi t j/n)\right)^2 \end{aligned}$$

22 Math

22.1 Gamma Function

- Ordinary: $\Gamma(s) = \int_0^\infty t^{s-1} e^{-t} dt$
- Upper incomplete: $\Gamma(s,x) = \int_x^\infty t^{s-1} e^{-t} dt$
- Lower incomplete: $\gamma(s,x) = \int_0^x t^{s-1} e^{-t} dt$
- $\Gamma(\alpha+1) = \alpha\Gamma(\alpha) \qquad \alpha > 1$
- $\Gamma(n) = (n-1)! \qquad n \in \mathbb{N}$
- $\Gamma(0) = \Gamma(-1) = \infty$
- $\Gamma(1/2) = \sqrt{\pi}$
- $\Gamma(-1/2) = -2\Gamma(1/2)$

22.2 Beta Function

- Ordinary: $\mathrm{B}(x,y) = \mathrm{B}(y,x) = \int_0^1 t^{x-1}(1-t)^{y-1}\,dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$
- Incomplete: $\mathrm{B}(x;\,a,b) = \int_0^x t^{a-1}(1-t)^{b-1}\,dt$
- Regularized incomplete:
  $$I_x(a,b) = \frac{\mathrm{B}(x;\,a,b)}{\mathrm{B}(a,b)} \overset{a,b\in\mathbb{N}}{=} \sum_{j=a}^{a+b-1} \frac{(a+b-1)!}{j!(a+b-1-j)!} x^j (1-x)^{a+b-1-j}$$
- $I_0(a,b) = 0 \qquad I_1(a,b) = 1$
- $I_x(a,b) = 1 - I_{1-x}(b,a)$

22.3 Series

Finite

- $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$
- $\sum_{k=1}^{n} (2k-1) = n^2$
- $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$
- $\sum_{k=1}^{n} k^3 = \left(\frac{n(n+1)}{2}\right)^2$
- $\sum_{k=0}^{n} c^k = \frac{c^{n+1}-1}{c-1} \qquad c \neq 1$

Binomial

- $\sum_{k=0}^{n} \binom{n}{k} = 2^n$
- $\sum_{k=0}^{n} \binom{r+k}{k} = \binom{r+n+1}{n}$
- $\sum_{k=0}^{n} \binom{k}{m} = \binom{n+1}{m+1}$
- Vandermonde's Identity:
  $\sum_{k=0}^{r} \binom{m}{k}\binom{n}{r-k} = \binom{m+n}{r}$
- Binomial Theorem:
  $\sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k = (a+b)^n$

Infinite

- $\sum_{k=0}^{\infty} p^k = \frac{1}{1-p}, \quad \sum_{k=1}^{\infty} p^k = \frac{p}{1-p} \qquad |p| < 1$
- $\sum_{k=0}^{\infty} kp^{k-1} = \frac{d}{dp}\left(\sum_{k=0}^{\infty} p^k\right) = \frac{d}{dp}\left(\frac{1}{1-p}\right) = \frac{1}{(1-p)^2} \qquad |p| < 1$
- $\sum_{k=0}^{\infty} \binom{r+k-1}{k} x^k = (1-x)^{-r} \qquad r \in \mathbb{N}^+$
- $\sum_{k=0}^{\infty} \binom{\alpha}{k} p^k = (1+p)^\alpha \qquad |p| < 1\,,\ \alpha \in \mathbb{C}$