UNIVERSITY OF CAMBRIDGE

MATHEMATICAL TRIPOS Part III

Friday, 3 June, 2016 9:00 am to 12:00 pm

PAPER 208

TIME SERIES AND MONTE CARLO INFERENCE

*Attempt no more than **FOUR** questions.*

*There are **SIX** questions in total.*

The questions carry equal weight.

STATIONERY REQUIREMENTS
Cover sheet
Treasury Tag
Script paper

SPECIAL REQUIREMENTS
None

You may not start to read the questions printed on the subsequent pages until instructed to do so by the Invigilator.

1

Let $\{\eta_t\}_{t\in\mathbb{Z}}$ be a strong White Noise with variance 1 and $\{X_t\}_{t\in\mathbb{Z}}$ be the process satisfying $\forall t \in \mathbb{Z}$:

$$X_t - \frac{1}{4}X_{t-1} = \eta_t - 3\eta_{t-1}$$

1. Give the definition of weak and strong White Noise.
2. Give the spectral density of X.
3. Is $\{\eta_t\}_{t\in\mathbb{Z}}$ the innovation process of $\{X_t\}_{t\in\mathbb{Z}}$?
4. Determine the ARMA equation relating $\{X_t\}_{t\in\mathbb{Z}}$ to its innovation process $\{\varepsilon_t\}_{t\in\mathbb{Z}}$.
5. Determine the $MA(\infty)$ form of X.
6. Compute its autocovariance function $\gamma_X(h)\,,\ h \in \mathbb{Z}$.
7. Compute the autocovariance function $\gamma_Z(h)\,,\ h \in \mathbb{Z}$ of the process $\{Z_t\}_{t\in\mathbb{Z}}$ defined $\forall t \in \mathbb{Z}$ as follows:

$$Z_t = a + bX_t + cX_t^2,$$

where a, b, c are some constants.
Hint: the following property of the Hermite polynomials might help.

The Hermite polynomial family $H_n(x)$ is defined for non negative integers n by:

$$H_n(x) = (-1)^n \exp\left(\frac{x^2}{2}\right)\left(\frac{d}{dx}\right)^n \left[\exp\left(-\frac{x^2}{2}\right)\right].$$

One can show (**not here !**) that if U and V are two **standard Gaussian random variables with correlation** ρ, then

$$\mathbb{E}\left[H_n(U)\,H_m(V)\right] = \begin{cases} 0, & \text{if } n \neq m \\ n!\rho^n, & \text{if } n = m. \end{cases}$$

2

1. The Figure 1 below shows a time series plot (a realization from a stationary process) and its associated (sample/empirical) AutoCorrelation Function (ACF) and Partial ACF. How do you select a model order from ACF and PACF ? What would be your modeling advices in this case? Justify your answer. Could you get any additional information out of these ACF and PACF?

2. What is the periodogram ? Give its formula when computed on T observations $X_1, \ldots, X_T$ of a zero mean stationary process $\{X_t\}$? Is it a good estimator (justify your answer)?

3. We have at our disposal a sample of T observations $X_1, \ldots, X_T$ from a stationary process $\{X_t\}_{t\in\mathbb{Z}}$ with mean μ.

   - Assume that its autocovariance function $\gamma_X(h)$ satisfies $\sum_{h\in\mathbb{Z}} |\gamma_X(h)| < \infty$. Prove that $T\mathbb{V}ar\left(\bar{X}_T\right) \to \sum_{h\in\mathbb{Z}} |\gamma_X(h)|$, where $\bar{X}_T = \frac{1}{T}\sum_{t=1}^{T} X_t$ is the empirical mean.
   - Under extra conditions one can show that the Central Limit Theorem applies. What can you say about the validity of the CLT when dealing with correlated data?
   - Now the process X is given. It is a stationary and causal AR(1) process defined $\forall t \in \mathbb{Z}$:

$$X_t = m + \phi X_{t-1} + \varepsilon_t,\ \varepsilon_t \overbrace{\sim}^{iid} \mathcal{N}\left(0, \sigma_\varepsilon^2\right)$$

     - Compute its mean $\mu = \mathbb{E}\left(X_t\right)$
     - What is the conditional law of $X_t\,|X_{t-1} = x_{t-1}$?
     - Compute the **conditional** Maximum likelihood estimator $\hat{\mu}$ of μ; is it unbiased ? Find its variance and check that it is consistent.

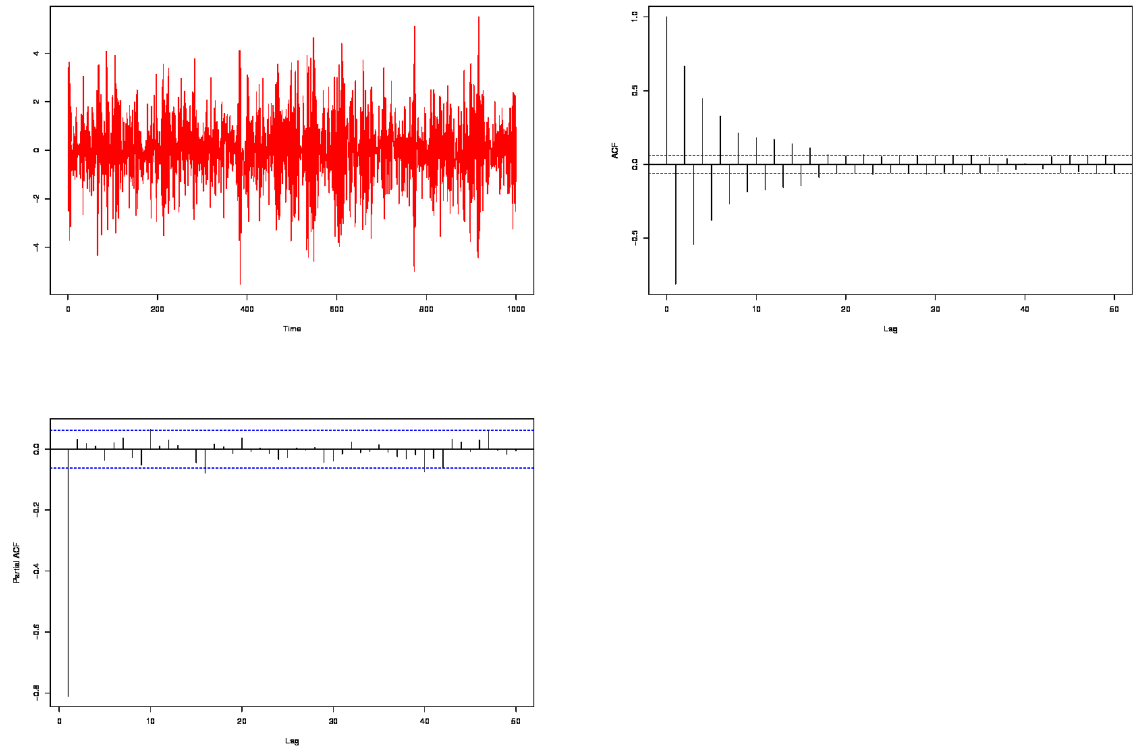

Figure 1

3

1. Figure 2 wrongly associates 4 time series plots (on the left) to 4 characteristic functions (on the right). The characteristic functions are either autocorrelation functions or spectral densities. Connect every time series to their characteristic function (justify your answer).

2. How would you define non stationary processes (in words)? Which of the previous models exhibit such non stationary behavior? Give 3 ways that are used to deal with some nonstationary behavior.

3. Consider the following processes defined $\forall t \in \mathbb{Z}$

$$
\begin{aligned}
Y_t &= \cos\left(\frac{2\pi t}{S}\right) + \varepsilon_t \\
Z_t &= \cos\left(\frac{2\pi t}{S}\right) \varepsilon_t
\end{aligned}
$$

where $\varepsilon_t \overbrace{\sim}^{iid} \mathcal{N}\left(0, \sigma^2\right)$ and $S \in \mathbb{N}^*$.

   - Are these processes stationary? If not, describe the nature of their nonstationarity.
   - Prove that, the application of an appropriate differentiation operation to $\{Y_t\}$ yields a stationary process.
   - Apply the same differentiation operator to $\{Z_t\}$. What are the consequences? Comment on this result from a practical point of view.
   - Show that $\{Z_t\}$ is periodically correlated.

4. We now turn our attention to periodically correlated processes and the associated Periodic ARMA models. In this question we consider a causal centered Periodic AR(1) of period T:

$$
X_{nT+\nu} = \phi(\nu) X_{nT+\nu-1} + \sigma(\nu) \varepsilon_{nT+\nu}
$$

where $\nu = 1, \ldots, T$ is the season and $n = 0, \ldots, N$ is the period. We define the periodic autocovariance function $\gamma_\nu(h) = \mathbb{E}\left[X_{nT+\nu} X_{nT+\nu-h}\right]$. We can estimate it using a sample version that we just write $\hat{\gamma}_\nu(h)$. Note also that $\gamma_\nu(h) = \gamma_{nT+\nu}(h)$ and $\gamma_\nu(-h) = \gamma_{\nu+h}(h)$:

   - Find the Yule-Walker equations.
   - From the previous part deduce an estimator of $\phi(\nu)$ and of $\sigma(\nu)$ for $\nu = 1, \ldots, T$.

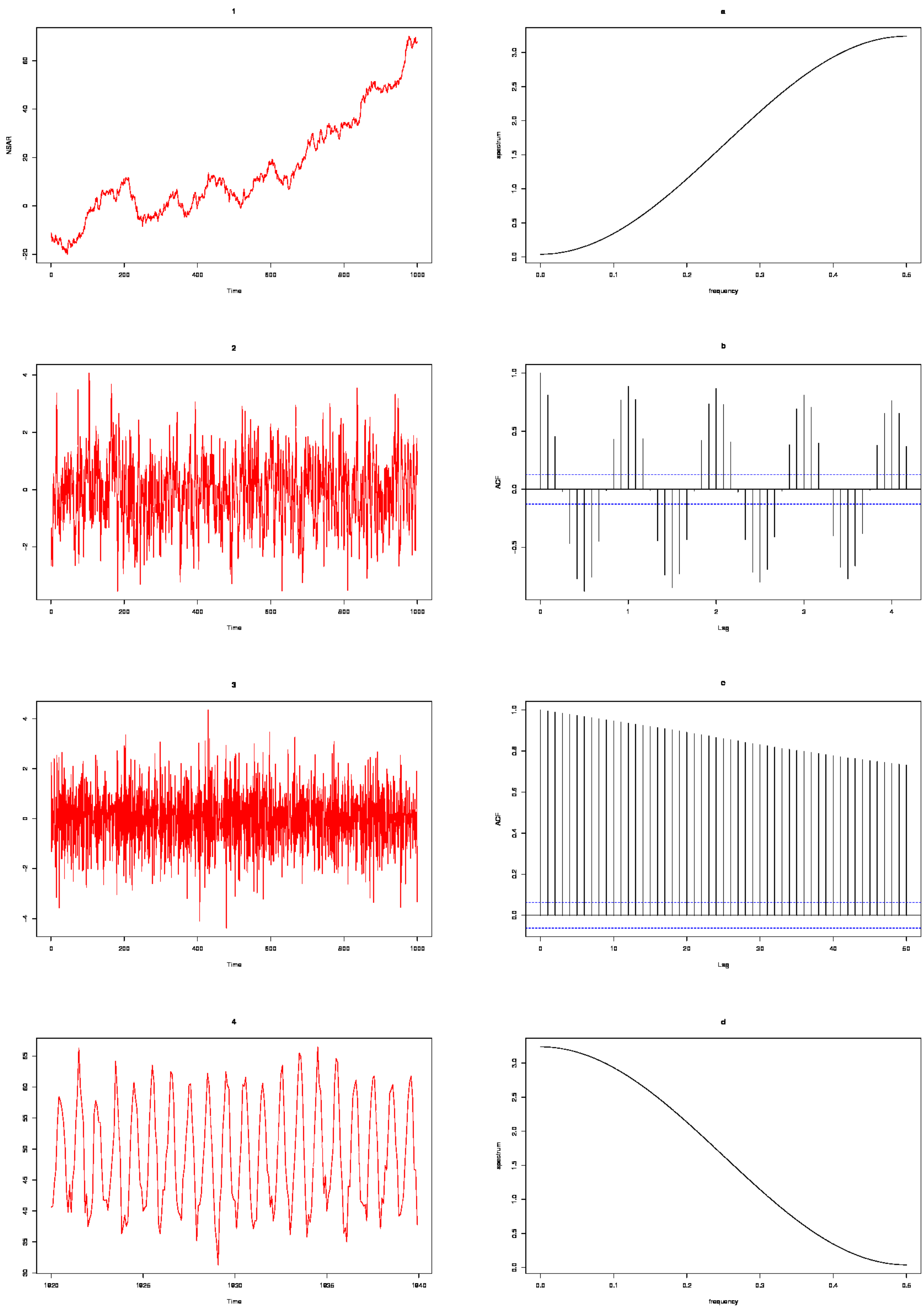

Figure 2: Time Series data on the left; ACF or spectrum on the right. Which characteristic function for which time series ?

4

(a) Present the Box-Muller algorithm to generate $X_1, X_2 \overset{iid}{\sim} N(0,1)$ from $U_1, U_2 \overset{iid}{\sim} Unif[0,1]$, and prove that it works.

(b) We want to estimate the mean of the truncated normal distribution (with density $f(x) = \frac{\phi(x)}{\int_0^1 \phi(x)dx}$, where $\phi()$ denotes the standard normal density). Apply the importance sampling with a right choice of importance distribution. What is the importance weight in this case? Now alternatively give an estimate of the mean by simulating from the standard normal distribution and keeping the draws in $[0,1]$. Explain why importance sampling works better.

(c) Given $X_1, \ldots, X_d \overset{iid}{\sim} N(0,1)$, and $\mu \in \mathbb{R}_d$ and Σ a $d \times d$ non-negative symmetric matrix (i.e. $\Sigma = \Sigma^T$ and $x^T \Sigma x \geqslant 0, \forall x \in \mathbb{R}_d$), how can one generate $X \sim N_d(\mu, \Sigma)$ (i.e. from a multivariate normal distribution)? Prove that your method works.

(d) Let $X_1, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$. Consider the following Bayesian model:

$$f(\mu, \sigma^2) = \frac{1}{\sigma^2}; \qquad (X_i|\mu, \sigma) \overset{iid}{\sim} N(\mu, \sigma^2), \quad i = 1, \ldots, n.$$

Notice that the prior for (μ, σ^2) is improper. Find an improper posterior PDF for this model. Provide the Gibbs sampler to sample from the posterior distribution by finding the suitable conditional PDFs. [Hint: Remember that the PDF of $Gamma(\alpha, \beta)$ is proportional to $y^{\alpha-1}e^{-\beta y}$.]

5

(a) Let $f(x)$ be a density function on $\mathbb{R}$, and $h : \mathbb{R} \to \mathbb{R}$ such that $\int_{\mathbb{R}} h(x)^2 f(x)dx < \infty$. Apply the central limit theorem in order to find the $(1-\alpha)\%$ confidence interval for the classical Monte Carlo estimator of the integral $\int_{\mathbb{R}} h(x)f(x)dx$ when the sample size is sufficiently large.

(b) Consider N samples produced by the accept-reject method on (f, g), where f, g are densities on $\mathbb{R}^d$ and $f \leqslant Mg$. Suppose that out of these samples, $(Z_1, \ldots, Z_n)$, $0 < n < N$, are the rejected subsamples. Show that

$$\frac{1}{n}\sum_{i=1}^{n} h(Z_i)\frac{(M-1)f(Z_i)}{Mg(Z_i) - f(Z_i)}$$

is an unbiased estimator of $E_f[h(x)]$, where $h : \mathbb{R}^d \to \mathbb{R}$.

(c) We apply the accept-reject method for generating from the positive half of a standard normal distribution ($f(x) = \frac{2}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, $x > 0$) using an exponential distribution (with improper density $g(x) = e^{-x}$). Show that with the right choice of M, on average, more than 75% of trials are accepted. [You do not need to prove your assertion about the average probability of acceptance.] [$\sqrt{\frac{\pi}{2e}} \approx 0.76$.]

(d) We want to estimate $l = P(S_R \leqslant x)$, where $S_R = \sum_{i=1}^{R} X_i$, R is a random variable with a given distribution (which is easy to sample from), and $X_i \overset{i.i.d}{\sim} F$ (which is also easy to sample from), and X_i are independent of R. Observe that for a fixed $R = r$,

$$F^r(x) := P(\sum_{i=1}^{r} X_i \leqslant x) = F(x - \sum_{i=2}^{r} X_i).$$

Use this in order to find a Monte Carlo estimator of l. Explain the steps you take in order to compute this estimator. [Hint: Use the law of total expectation for l.]

6

(a) Define the Metropolis-Hastings algorithm for obtaining samples from a (possibly improper) density $\pi(x)$. Define also what it means for a transition kernel density K and a probability measure π to be in detailed balance.

(b) Prove that the transition kernel density of the Markov chain generated by Metropolis-Hastings and π are in detailed balance.

(c) Show that the transition kernel of the Gibbs sampler (for a density π on $\mathbb{R}^p$) is the composition of p Metropolis-Hastings transition kernels, each of which have acceptance probability 1.

(d) Let the random vector $X = (X_1, X_2)$ have the following two-dimensional PDF:

$$f(x) = c\exp(-(x_1^2 x_2^2 + x_1^2 + x_2^2 - 8x_1 - 8x_2)/2),$$

where c is the normalizing constant. Suppose we wish to estimate $E[X_1]$ via the classical Monte Carlo estimator $\frac{1}{N}\sum_{t=1}^{N} X_{t,1}$. We want to sample $\{X_t\}$ using random walk Metropolis-Hastings. Which theorem should we use in order to ensure that the Monte Carlo estimator is consistent? [You do not need to state the regularity conditions explicitly.] Provide the exact procedure for random walk Metropolis-Hastings sampling from f assuming you can sample from standard normal and uniform distributions. [No need to prove anything.] Discuss how changing the variance of your proposal distribution affects the generated chain.

END OF PAPER