

# Supplementary Material for Quantifying the Trendiness of Trends

Andreas Kryger Jensen and Claus Thorn Ekstrøm  
Biostatistics, Institute of Public Health, University of Copenhagen  
aeje@sund.ku.dk, ekstrom@sund.ku.dk

03 October, 2020

## A Proof of Proposition 1

Let  $\mathbf{Y} = (Y_1, \dots, Y_n)$  and  $\mathbf{t} = (t_1, \dots, t_n)$  be the vectors of observed outcomes and associated sampling times. From the data generating model we observe that the marginal distribution of the vector of observed outcomes  $\mathbf{Y} \mid \mathbf{t}, \Theta$  is

$$\begin{aligned} P(\mathbf{Y} \mid \mathbf{t}, \Theta) &= \int P(\mathbf{Y} \mid f(\mathbf{t}), \mathbf{t}, \Theta) dP(f(\mathbf{t}) \mid \mathbf{t}, \Theta) \\ &= N(\mu_\beta(\mathbf{t}), C_\theta(\mathbf{t}, \mathbf{t}) + \sigma^2 I) \end{aligned}$$

where  $\mu_\beta(\mathbf{t}) = (\mu_\beta(t_1), \dots, \mu_\beta(t_n))$ ,  $C_\theta(\mathbf{t}, \mathbf{t})$  is the  $n \times n$  covariance matrix obtained by evaluating  $C_\theta(s, t)$  at  $\{(s, t) \in \mathbf{t} \times \mathbf{t}\}$  and  $I$  is an  $n \times n$  identity matrix. This implies that the joint distribution of  $\mathbf{Y}$  and the latent functions  $(f, df, d^2f)$  evaluated at an arbitrary vector of time points  $\mathbf{t}^*$  is

$$\begin{bmatrix} f(\mathbf{t}^*) \\ df(\mathbf{t}^*) \\ d^2f(\mathbf{t}^*) \\ \mathbf{Y} \end{bmatrix} \mid \mathbf{t}, \Theta \sim N \left( \begin{bmatrix} \mu_\beta(\mathbf{t}^*) \\ d\mu_\beta(\mathbf{t}^*) \\ d^2\mu_\beta(\mathbf{t}^*) \\ \mu_\beta(\mathbf{t}) \end{bmatrix}, \begin{bmatrix} C_\theta(\mathbf{t}^*, \mathbf{t}^*) & \partial_2 C_\theta(\mathbf{t}^*, \mathbf{t}^*) & \partial_2^2 C_\theta(\mathbf{t}^*, \mathbf{t}^*) & C_\theta(\mathbf{t}^*, \mathbf{t}) \\ \partial_1 C_\theta(\mathbf{t}^*, \mathbf{t}^*) & \partial_1 \partial_2 C_\theta(\mathbf{t}^*, \mathbf{t}^*) & \partial_1 \partial_2^2 C_\theta(\mathbf{t}^*, \mathbf{t}^*) & \partial_1 C_\theta(\mathbf{t}^*, \mathbf{t}) \\ \partial_1^2 C_\theta(\mathbf{t}^*, \mathbf{t}^*) & \partial_1^2 \partial_2 C_\theta(\mathbf{t}^*, \mathbf{t}^*) & \partial_1^2 \partial_2^2 C_\theta(\mathbf{t}^*, \mathbf{t}^*) & \partial_1^2 C_\theta(\mathbf{t}^*, \mathbf{t}) \\ C_\theta(\mathbf{t}, \mathbf{t}^*) & \partial_2 C_\theta(\mathbf{t}, \mathbf{t}^*) & \partial_2^2 C_\theta(\mathbf{t}, \mathbf{t}^*) & C_\theta(\mathbf{t}, \mathbf{t}) + \sigma^2 I \end{bmatrix} \right)$$

where  $\partial_j^k$  denotes the  $k$ 'th order partial derivative with respect to the  $j$ 'th variable.

By the standard formula for deriving conditional distributions in a multivariate normal model, the posterior distribution of  $(f, df, d^2f)$  evaluated at the  $p$  time points in  $\mathbf{t}^*$  is

$$\begin{bmatrix} f(\mathbf{t}^*) \\ df(\mathbf{t}^*) \\ d^2f(\mathbf{t}^*) \end{bmatrix} \mid \mathbf{Y}, \mathbf{t}, \Theta \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

where  $\boldsymbol{\mu} \in \mathbb{R}^{3p}$  is the column vector of posterior expectations and  $\boldsymbol{\Sigma} \in \mathbb{R}^{3p \times 3p}$  is the joint posterior covariance matrix, and these are given by

$$\begin{aligned} \boldsymbol{\mu} &= \begin{bmatrix} \mu_\beta(\mathbf{t}^*) \\ d\mu_\beta(\mathbf{t}^*) \\ d^2\mu_\beta(\mathbf{t}^*) \end{bmatrix} + \begin{bmatrix} C_\theta(\mathbf{t}^*, \mathbf{t}) \\ \partial_1 C_\theta(\mathbf{t}^*, \mathbf{t}) \\ \partial_1^2 C_\theta(\mathbf{t}^*, \mathbf{t}) \end{bmatrix} K_{\theta, \sigma}(\mathbf{t}, \mathbf{t})^{-1} (\mathbf{Y} - \mu_\beta(\mathbf{t})) \\ \boldsymbol{\Sigma} &= \begin{bmatrix} C_\theta(\mathbf{t}^*, \mathbf{t}^*) & \partial_2 C_\theta(\mathbf{t}^*, \mathbf{t}^*) & \partial_2^2 C_\theta(\mathbf{t}^*, \mathbf{t}^*) \\ \partial_1 C_\theta(\mathbf{t}^*, \mathbf{t}^*) & \partial_1 \partial_2 C_\theta(\mathbf{t}^*, \mathbf{t}^*) & \partial_1 \partial_2^2 C_\theta(\mathbf{t}^*, \mathbf{t}^*) \\ \partial_1^2 C_\theta(\mathbf{t}^*, \mathbf{t}^*) & \partial_1^2 \partial_2 C_\theta(\mathbf{t}^*, \mathbf{t}^*) & \partial_1^2 \partial_2^2 C_\theta(\mathbf{t}^*, \mathbf{t}^*) \end{bmatrix} - \begin{bmatrix} C_\theta(\mathbf{t}^*, \mathbf{t}) \\ \partial_1 C_\theta(\mathbf{t}^*, \mathbf{t}) \\ \partial_1^2 C_\theta(\mathbf{t}^*, \mathbf{t}) \end{bmatrix} K_{\theta, \sigma}(\mathbf{t}, \mathbf{t})^{-1} \begin{bmatrix} C_\theta(\mathbf{t}, \mathbf{t}^*) \\ \partial_2 C_\theta(\mathbf{t}, \mathbf{t}^*) \\ \partial_2^2 C_\theta(\mathbf{t}, \mathbf{t}^*) \end{bmatrix}^T \end{aligned}$$

where  $K_{\theta,\sigma}(\mathbf{t}, \mathbf{t}) = C_{\theta}(\mathbf{t}, \mathbf{t}) + \sigma^2 I$ . Partitioning  $\boldsymbol{\mu}$  and  $\boldsymbol{\Sigma}$  as

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_f(\mathbf{t}^* | \boldsymbol{\Theta}) \\ \mu_{df}(\mathbf{t}^* | \boldsymbol{\Theta}) \\ \mu_{d^2f}(\mathbf{t}^* | \boldsymbol{\Theta}) \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} \Sigma_f(\mathbf{t}^*, \mathbf{t}^* | \boldsymbol{\Theta}) & \Sigma_{f,df}(\mathbf{t}^*, \mathbf{t}^* | \boldsymbol{\Theta}) & \Sigma_{f,d^2f}(\mathbf{t}^*, \mathbf{t}^* | \boldsymbol{\Theta}) \\ \Sigma_{f,df}(\mathbf{t}^*, \mathbf{t}^* | \boldsymbol{\Theta}) & \Sigma_{df}(\mathbf{t}^*, \mathbf{t}^* | \boldsymbol{\Theta}) & \Sigma_{df,d^2f}(\mathbf{t}^*, \mathbf{t}^* | \boldsymbol{\Theta}) \\ \Sigma_{d^2f,f}(\mathbf{t}^*, \mathbf{t}^* | \boldsymbol{\Theta}) & \Sigma_{d^2f,df}(\mathbf{t}^*, \mathbf{t}^* | \boldsymbol{\Theta}) & \Sigma_{d^2f}(\mathbf{t}^*, \mathbf{t}^* | \boldsymbol{\Theta}) \end{bmatrix}$$

and completing the matrix algebra, we obtain the expressions of the individual components given in the Proposition.

## B Proof of Proposition 3

Rice showed in section 3.3. of Rice (1945) that the expected number of zero-crossings of a Gaussian process  $X$  on an interval  $\mathcal{I}$  is given by

$$\int_{\mathcal{I}} \int_{-\infty}^{\infty} |v| f_{X(t), dX(t)}(0, v) dv dt \quad (1)$$

where  $f_{X(t), dX(t)}$  is the joint density function of  $X$  and its derivative  $dX$  at time  $t$ . To derive the expression for the Expected Trend Instability we must apply the Rice formula to the joint posterior distribution of  $(df, d^2f)$ . From Proposition 1 the distribution of  $(df, d^2f) | \mathbf{Y}, \mathbf{t}, \boldsymbol{\Theta}$  is bivariate normal for each  $t$ .

Let  $\mu_{df}$ ,  $\mu_{d^2f}$ ,  $\Sigma_{df}$  and  $\Sigma_{d^2f}$  be defined as in Proposition 1 and define further

$$\omega(t | \boldsymbol{\Theta}) = \frac{\Sigma_{df, d^2f}(t, t | \boldsymbol{\Theta})}{\Sigma_{df}(t, t | \boldsymbol{\Theta})^{1/2} \Sigma_{d^2f}(t, t | \boldsymbol{\Theta})^{1/2}}$$

as the posterior point-wise cross-correlation function between  $df$  and  $d^2f$ . The joint posterior density function of  $(df, d^2f)$  at any time  $t$  evaluated at  $(0, v)$  can be factorized as

$$f_{df(t), d^2f(t)}(0, v) = c_1(t) e^{c_2(t)} e^{-c_3(t)v^2 - 2c_4(t)v}$$

where  $c_1, \dots, c_4$  are functions of time given by

$$\begin{aligned} c_1(t) &= (2\pi)^{-1} \Sigma_{df}(t, t | \boldsymbol{\Theta})^{-1/2} \Sigma_{d^2f}(t, t | \boldsymbol{\Theta})^{-1/2} (1 - \omega(t | \boldsymbol{\Theta}))^{-1/2} \\ c_2(t) &= \frac{\mu_{df}(t | \boldsymbol{\Theta})^2}{2\Sigma_{df}(t, t | \boldsymbol{\Theta})(\omega(t | \boldsymbol{\Theta})^2 - 1)} + \frac{\mu_{d^2f}(t | \boldsymbol{\Theta})^2}{2\Sigma_{d^2f}(t, t | \boldsymbol{\Theta})(\omega(t | \boldsymbol{\Theta})^2 - 1)} \\ &\quad - \frac{\mu_{df}(t | \boldsymbol{\Theta})\mu_{d^2f}(t | \boldsymbol{\Theta})\omega(t | \boldsymbol{\Theta})}{\Sigma_{df}(t, t | \boldsymbol{\Theta})^{1/2}\Sigma_{d^2f}(t, t | \boldsymbol{\Theta})^{1/2}(\omega(t | \boldsymbol{\Theta})^2 - 1)} \\ c_3(t) &= -\frac{1}{2}\Sigma_{d^2f}(t, t | \boldsymbol{\Theta})^{-1}(\omega(t | \boldsymbol{\Theta})^2 - 1)^{-1} \\ c_4(t) &= -\frac{\mu_{df}(t | \boldsymbol{\Theta})\Sigma_{d^2f}(t, t | \boldsymbol{\Theta})^{1/2}\omega(t | \boldsymbol{\Theta}) - \mu_{d^2f}(t | \boldsymbol{\Theta})\Sigma_{df}(t, t | \boldsymbol{\Theta})^{1/2}}{2\Sigma_{d^2f}(t, t | \boldsymbol{\Theta})(\omega(t | \boldsymbol{\Theta})^2 - 1)\Sigma_{df}(t, t | \boldsymbol{\Theta})^{1/2}} \end{aligned}$$

Let  $d\text{ETI}(t | \boldsymbol{\Theta})$  denote the inner integral in Equation (1). Using the factorization of the joint posterior density we may write it was

$$\begin{aligned} d\text{ETI}(t | \boldsymbol{\Theta}) &= \int_{-\infty}^{\infty} |v| f_{df(t), d^2f(t)}(0, v) dv \\ &= c_1(t) e^{c_2(t)} \int_{-\infty}^{\infty} |v| e^{-c_3(t)v^2 - 2c_4(t)v} dv \\ &= c_1(t) e^{c_2(t)} \left( \int_0^{\infty} v e^{-c_3(t)v^2 + 2c_4(t)v} dv + \int_0^{\infty} v e^{-c_3(t)v^2 - 2c_4(t)v} dv \right) \end{aligned} \quad (2)$$

Because  $c_3(t) > 0$  for all  $t$  since  $\Sigma_{d^2f}(t, t | \Theta) > 0$  and  $|\omega(t | \Theta)| < 1$  by Assumption A4 we obtain the following solution for the type of integral in the previous display by using formula 5 in section 3.462 on page 365 of Gradshteyn and Ryzhik (2014)

$$\int_0^\infty v e^{-c_3(t)v^2 \pm 2c_4(t)v} dv = \frac{1}{2c_3(t)} \pm \frac{c_4(t)}{2c_3(t)} \frac{\pi^{1/2}}{c_3(t)^{1/2}} e^{\frac{c_4(t)^2}{c_3(t)}} \left( 1 \pm \operatorname{Erf} \left( \frac{c_4(t)}{\sqrt{c_3(t)}} \right) \right) \quad (3)$$

where  $\operatorname{Erf}: x \mapsto 2\pi^{-1} \int_0^x e^{-u^2} du$  is the error function. Combining Equations (2) and (3) we may express  $d\text{ETI}$  as

$$d\text{ETI}(t | \Theta) = c_1(t) e^{c_2(t)} \left( \frac{1}{c_3(t)} + \frac{c_4(t)}{c_3(t)} \frac{\pi^{1/2}}{c_3(t)^{1/2}} e^{\frac{c_4(t)^2}{c_3(t)}} \operatorname{Erf} \left( \frac{c_4(t)}{\sqrt{c_3(t)}} \right) \right)$$

Defining  $\zeta(t | \Theta) = \sqrt{2}c_4(t)c_3(t)^{-1/2}$  and collecting some terms, the index can be rewritten as

$$d\text{ETI}(t | \Theta) = \frac{c_1(t)}{c_3(t)} \left( e^{c_2(t)} + \frac{\pi^{1/2}}{2^{1/2}} e^{\frac{c_4(t)^2}{c_3(t)} + c_2(t)} \zeta(t) \operatorname{Erf} \left( \frac{\zeta(t | \Theta)}{2^{1/2}} \right) \right)$$

Straightforward arithmetic calculations show that

$$\frac{c_4(t)^2}{c_3(t)} + c_2(t) = -\frac{\mu_{df}(t | \Theta)^2}{2\Sigma_{df}(t, t | \Theta)}, \quad c_2(t) = -\frac{1}{2} \left( \zeta(t | \Theta)^2 + \frac{\mu_{df}(t | \Theta)^2}{\Sigma_{df}(t, t | \Theta)} \right)$$

and by defining  $\phi: x \mapsto (2\pi)^{-1/2} e^{-x^2}$  as the density function of the standard normal distribution we may write  $e^{\frac{c_4(t)^2}{c_3(t)} + c_2(t)} = (2\pi)^{1/2} \phi \left( \frac{\mu_{df}(t | \Theta)}{\Sigma_{df}(t, t | \Theta)^{1/2}} \right)$  and  $e^{c_2(t)} = 2\pi \phi(\zeta(t)) \phi \left( \frac{\mu_{df}(t | \Theta)}{\Sigma_{df}(t, t | \Theta)^{1/2}} \right)$  which leads to

$$d\text{ETI}(t | \Theta) = \frac{c_1(t)}{c_3(t)} \pi \phi \left( \frac{\mu_{df}(t | \Theta)}{\Sigma_{df}(t, t | \Theta)^{1/2}} \right) \left( 2\phi(\zeta(t | \Theta)) + \zeta(t | \Theta) \operatorname{Erf} \left( \frac{\zeta(t | \Theta)}{2^{1/2}} \right) \right)$$

Standard arithmetics show that

$$\frac{c_1(t)}{c_3(t)} = \frac{1}{\pi} \frac{\Sigma_{d^2f}(t, t | \Theta)^{1/2}}{\Sigma_{df}(t, t | \Theta)^{1/2}} \left( 1 - \omega(t | \Theta)^2 \right)^{1/2}$$

and we finally obtain the expression

$$d\text{ETI}(t | \Theta) = \lambda(t | \Theta) \phi \left( \frac{\mu_{df}(t | \Theta)}{\Sigma_{df}(t, t | \Theta)^{1/2}} \right) \left( 2\phi(\zeta(t | \Theta)) + \zeta(t | \Theta) \operatorname{Erf} \left( \frac{\zeta(t | \Theta)}{2^{1/2}} \right) \right)$$

where  $\lambda$  and  $\zeta$  are given by

$$\begin{aligned} \lambda(t | \Theta) &= \frac{\Sigma_{d^2f}(t, t | \Theta)^{1/2}}{\Sigma_{df}(t, t | \Theta)^{1/2}} \left( 1 - \omega(t | \Theta)^2 \right)^{1/2} \\ \zeta(t | \Theta) &= \frac{\mu_{df}(t | \Theta) \Sigma_{d^2f}(t, t | \Theta)^{1/2} \omega(t) \Sigma_{df}(t, t | \Theta)^{-1/2} - \mu_{d^2f}(t | \Theta)}{\Sigma_{d^2f}(t, t | \Theta)^{1/2} \left( 1 - \omega(t | \Theta)^2 \right)^{1/2}} \end{aligned}$$

By definition

$$\text{ETI}(\mathcal{I} | \Theta) = \int_{\mathcal{I}} d\text{ETI}(t | \Theta) dt$$

which completes the proof.

## C Zero-crossings of $f$ and $df$ in the zero-mean stationary case

Let  $f \sim \mathcal{GP}(0, C_\theta(\cdot, \cdot))$  where the  $C_\theta$  is either the Squared Exponential or Rational Quadratic covariance function. We look at the expected number of zero-crossings on an interval by either  $f$  and  $df$  as given by the Rice formula in Equation (1) with either  $X(t) = f(t)$  or  $X(t) = df(t)$ . In this case the expressions simplifies immensely due to the zero means of both  $f$ ,  $df$ , and  $d^2f$  and because  $\text{Cov}[f(t), df(t)] = 0$  and  $\text{Cov}[df(t), d^2f(t)] = 0$ . The latter is a result of using a stationary covariance function for the prior distribution of  $f$  (Cramer and Leadbetter 1967). In this stationary case local expected number of zero-crossing of  $f$  and  $df$  are given by

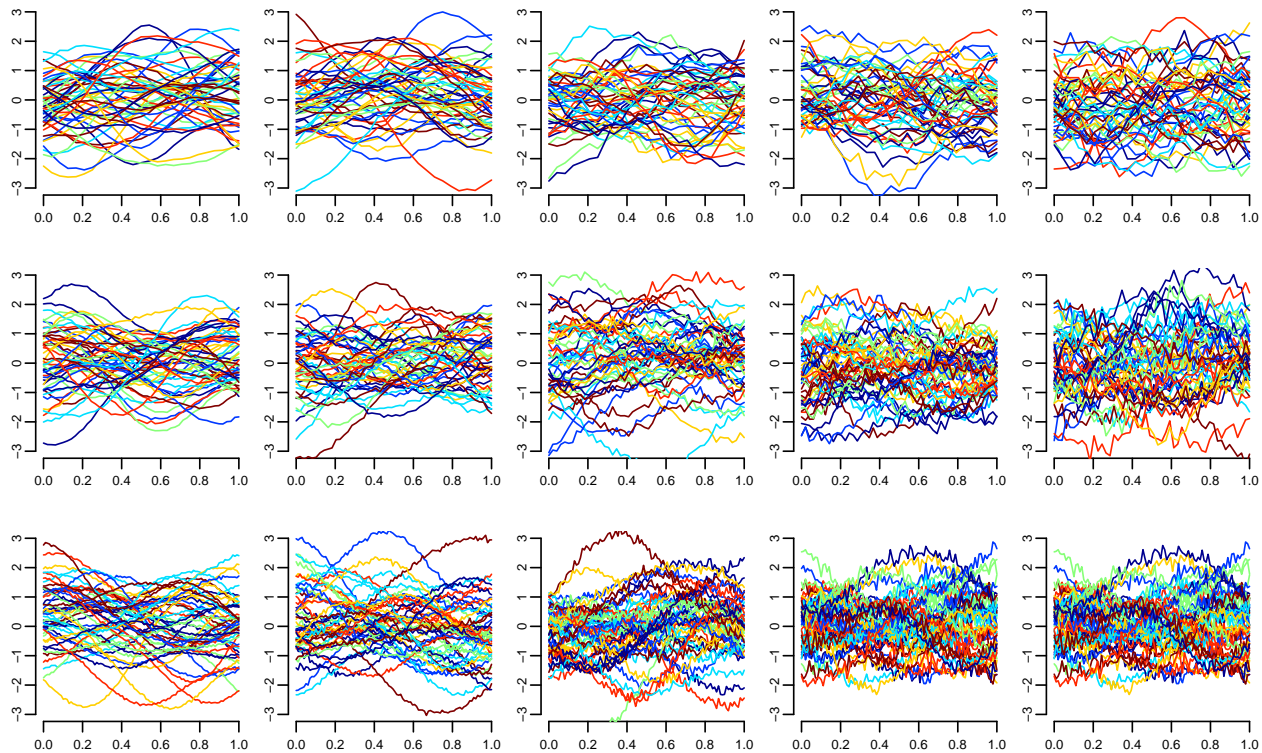
$$\frac{\partial_1 \partial_2 C_\theta(s, t) \Big|_{s=t}^{1/2}}{\pi C_\theta(t, t)^{1/2}} \quad \text{and} \quad \frac{\partial_1^2 \partial_2^2 C_\theta(s, t) \Big|_{s=t}^{1/2}}{\pi \partial_1 \partial_2 C_\theta(s, t) \Big|_{s=t}^{1/2}}$$

respectively. It then follows that

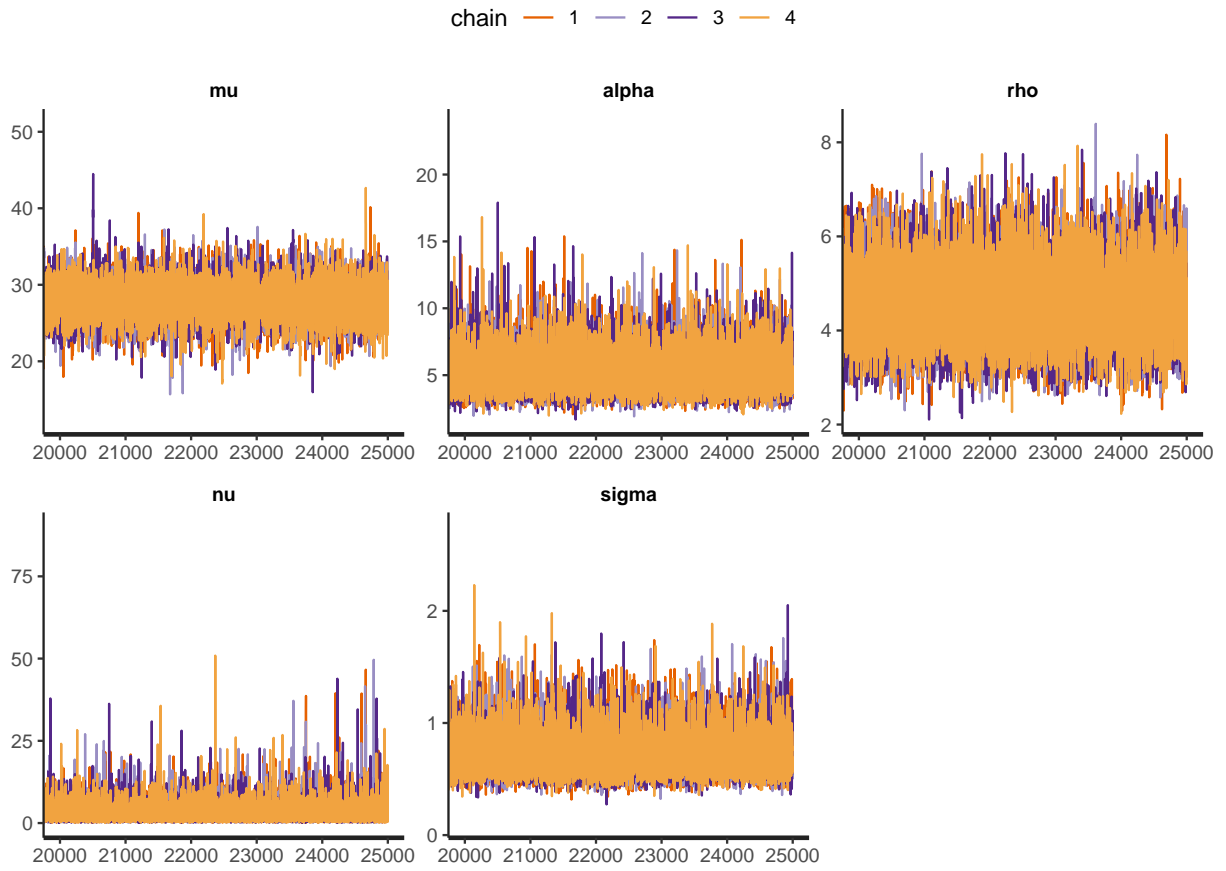
$$\begin{aligned} C_\theta^{\text{SE}}(t, t) = \sigma^2, \quad \partial_1 \partial_2 C_\theta^{\text{SE}}(s, t) \Big|_{s=t} &= \frac{\sigma^2}{\rho^2}, \quad \partial_1^2 \partial_2^2 C_\theta^{\text{SE}}(s, t) \Big|_{s=t} = \frac{3\sigma^2}{\rho^4} \\ C_\theta^{\text{RQ}}(t, t) = \sigma^2, \quad \partial_1 \partial_2 C_\theta^{\text{RQ}}(s, t) \Big|_{s=t} &= \frac{\sigma^2}{\rho^2}, \quad \partial_1^2 \partial_2^2 C_\theta^{\text{RQ}}(s, t) \Big|_{s=t} = \frac{2\sigma^2(1 + \nu)}{\nu \rho^4} \end{aligned}$$

and the local expected number of zero-crossings of  $f$  and  $df$  for either the Squared Exponential and the Rational Quadratic covariance functions are

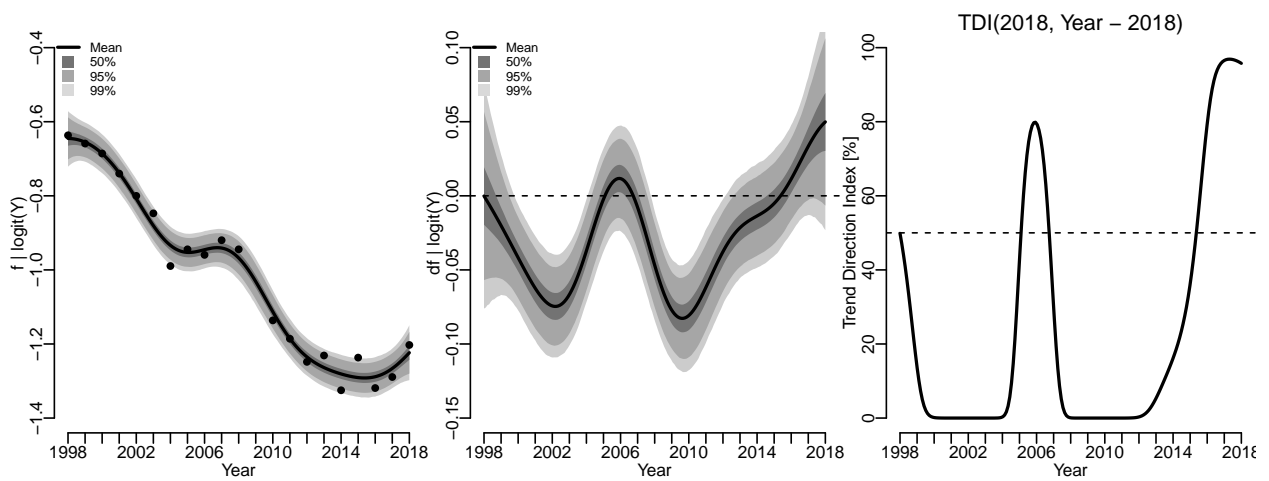
$$\begin{aligned} \frac{\partial_1 \partial_2 C_\theta^{\text{SE}}(s, t) \Big|_{s=t}^{1/2}}{\pi C_\theta^{\text{SE}}(t, t)^{1/2}} &= \frac{1}{\pi \rho}, & \frac{\partial_1^2 \partial_2^2 C_\theta^{\text{SE}}(s, t) \Big|_{s=t}^{1/2}}{\pi \partial_1 \partial_2 C_\theta^{\text{SE}}(s, t) \Big|_{s=t}^{1/2}} &= \frac{3^{1/2}}{\pi \rho} \\ \frac{\partial_1 \partial_2 C_\theta^{\text{RQ}}(s, t) \Big|_{s=t}^{1/2}}{\pi C_\theta^{\text{RQ}}(t, t)^{1/2}} &= \frac{1}{\pi \rho}, & \frac{\partial_1^2 \partial_2^2 C_\theta^{\text{RQ}}(s, t) \Big|_{s=t}^{1/2}}{\pi \partial_1 \partial_2 C_\theta^{\text{RQ}}(s, t) \Big|_{s=t}^{1/2}} &= \frac{3^{1/2}}{\pi \rho} (1 + \nu^{-1})^{1/2} \end{aligned}$$



**Figure 1:** 50 random sample paths from each simulation scenario.



**Figure 2:** Trace plots of the last 5,000 MCMC iterations for the hyper-parameters in the smoking application.



**Figure 3:** Trend analysis of smoking data with logit transformed outcome.

## Bibliography

Cramer, Harald, and M. R. Leadbetter. 1967. *Stationary and Related Stochastic Processes – Sample Function Properties and Their Applications*. John Wiley & Sons, Inc.

Gradshteyn, Izrail Solomonovich, and Iosif Moiseevich Ryzhik. 2014. *Table of Integrals, Series, and Products*. Academic Press.

Rice, Stephen O. 1945. “Mathematical Analysis of Random Noise, II.” *Bell System Technical Journal* 24 (1): 46–156.