Tests and confidence intervals for the location parameter in orthogonal $FEXP$ models

Jan Beran
Department of Mathematics and Statistics
University of Konstanz
78457 Konstanz, Germany

Abstract

Confidence intervals and tests for the location parameter are considered for time series generated by $FEXP$ models. Since these tests mainly depend on the unknown fractional differencing parameter d, the distribution of $\hat{d}$ plays a major role. An exact closed form expresssion for the asymptotic variance of $\hat{d}$ is given for $FEXP$ models with cosine functions. It is shown that the variance increases linearily with the order p of the model. An alternative $FEXP$ model with orthogonal components is proposed for which the asymptotic variance of $\hat{d}$ does not depend on p. Tables of quantiles of the test statistic are given for both model classes.,

Key words: t-test, long-range dependence, short-range dependence, antipersistence, location estimation, confidence interval, $FEXP$ model.

1 Introduction

Let X_t $(t = 1, 2, ...)$ be a second order stationary process with expected value μ, autocovariances $\gamma(k) = \text{cov}(X_t, X_{t+k})$ and spectral density $f = (2\pi)^{-1} \sum_{k=-\infty}^{\infty} \gamma(k) \exp(ik\lambda)$, $\lambda \in [-\pi, \pi]$. Assume that f is continuous in $[-\pi, 0) \cup (0, \pi]$ and, as $\lambda \to 0$,

$$f(\lambda) \sim c_f |\lambda|^{-2d} \tag{1}$$

for some $-\frac{1}{2} < d < \frac{1}{2}$ and $0 < c_f < \infty$. Here " $\sim$ means that the left divided by the right hand side converges to one. Consider the problem of constructing $100(1-\alpha)\%$ confidence intervals for the expected value $\mu = E(X_i)$, or equivalently, testing $H_o : \mu = \mu_o$ at a level of significance $\alpha \in (0, 1)$. It is well known that standard tests and confidence intervals based on the t-statistic $T^o = \sqrt{n}(\bar{x} - \mu_o)/s$, with $\bar{x} = n^{-1} \sum x_i$ and $s^2 = (n-1)^{-1} \sum (x_i - \bar{x})^2$, and standard normal or $t_{n-1}-$quantiles are unreliable in the presence of dependence, in particular if the autocorrelations are not summable (see e.g. Mandelbrot and Wallis 1969, Beran 1989, 1994) . Asymptotically, the rate at which $\text{var}(\bar{x})$ decays to zero depends on the behaviour of the spectral density at the origin. Three cases can be distinguished within the framework given by (1): Short-range dependence with $d = 0$, f everywhere bounded and continuous in $[-\pi, \pi]$, and $\lim_{\lambda \to 0} f(\lambda) = c_f \in (0, \infty)$; long-range dependence with $d > 0$, f diverging to infinity at zero; and antipersistence with $d < 0$, and $f(0) = 0$. The variance of the sample mean is proportional to n^{2d-1}. More specifically, we have (see e.g. Adenstedt 1974, Samarov and Taqqu 1988, Beran 1994):

Proposition 1 *Let*

$$\nu(d) = \frac{2\Gamma(1-2d)\sin(\pi d)}{d(2d+1)} \tag{2}$$

with $\nu(0) = \lim_{d\to 0} \nu(d) = 2\pi$. *Then, under the assumptions above*

$$v = \mathrm{var}(\bar{x}) = n^{-1}\nu(d)f(\frac{1}{n}) + o(n^{2d-1}) = n^{2d-1}\nu(d)c_f + o(n^{2d-1}). \tag{3}$$

Thus, the usual n^{-1} rate of convergence is achieved for $d = 0$ only, whereas the rate is slower for long-range dependence and faster under antipersistence. As a result, confidence intervals based on T^o and the standard normal distribution are too small (with an asymptotic coverage probability of zero) under long-range dependence, whereas they are unnecessarily large (with an asymptotic coverage probability of one) under antipersistence.

Beran (1989) proposed a modified t-test that is valid under long memory and models in the neighbourhood of fractional Gaussian noise. More generally, the statistic in Beran (1989) can be adapted to any parametric class of models $f(\lambda) = f(\lambda;\theta)$ (see Beran 1994, chapter 8), such as fractional ARIMA (Hosking 1981, Granger and Joyeux 1980) or fractional exponential models (Beran 1993), in combination with a consistent model choice criterion (see e.g. Beran et al. 1999). The distribution of $T = (\bar{x} - \mu)/\sqrt{v(\hat{\theta})}$ can be approximated by the distribution of $Y = Z_1 n^{Z_2\sqrt{w/n}}$ where Z_1, Z_2 are independent standard normal random variables and w is the asymptotic variance of $\sqrt{n}(\hat{d} - d)$ (see Beran 1994). Quantiles of Y can be obtained by simulations. However, in general, w and thus the quantiles of T depend on $\hat{\theta}$ so that simulations need to be done afresh for each data set. This is not the case for $FEXP$ models, since there the asymptotic distribution of $\hat{d}$ only depends on the order p of the model.

In this note, we exploit this property and obtain a simple testing procedure by considering two $FEXP$ models based on orthogonal functions. Orthogonality makes it possible to give closed form formulas for the asymptotic variance of $\hat{d}$. Approximate distribution free quantiles of the test statistic can then be given and tabulated as a function of n and p or even as a function of n only. The method is valid under short-memory, long-memory and antipersistence.

2 $FEXP$ models

$FEXP$ models were introduced in Beran (1993) as a generalization of exponential models by Bloomfield (1973; also see Diggle 1990). An $FEXP(p)$ model is a second order stationary process with spectral density

$$f(\lambda) = \exp(\sum_{j=0}^{p+1} \beta_j \log g_j(\lambda)) \tag{4}$$

where $p \geq 0$ is an integer, $\beta = (\beta_o, \beta_1, ..., \beta_{p+1}) \in R^{p+2}$, $-1 < \beta_1 < 1$, $g_o(\lambda) = 1$, $g_1(\lambda)/\log|\lambda| \to 1$ (as $|\lambda| \to 0$), and $g_j(.) \in C[-\pi,\pi]$ ($j = 2, ..., p+1$). Here, the unknown parameter vector $\theta = (\theta_1, ..., \theta_{p+2})$ is equal to $\beta = (\beta_o, ..., \beta_{p+1})$. The interpretation of the parameters is a follows: β_o is the scale parameter; $\beta_1 = -2d$ models the long memory behaviour ($\beta = 0$ for short memory; $0 < \beta < 1$ for antipersistence; $-1 < \beta < 0$ for long memory); β_j ($j \geq 2$) are parameters that allow for flexible modelling of short-range dependence. A typical choice for g_1 is $g_1(\lambda) = \log|1 - \exp(i\lambda)|$. In this case, the spectral density of the $FEXP(0)$ model is identical with the spectral density of a fractional ARIMA(0,d,0) process (see Granger and Joyeux 1980, Hosking 1981). A typical choice for g_j ($j \geq 2$) is $g_j(\lambda) = \cos\{(j-1)\lambda\}$. In the following an $FEXP$ model with $g_1(\lambda) = \log|1-\exp(i\lambda)|$ and $g_j(\lambda) = \cos\{(j-1)\lambda\}$ ($j \geq 2$) will be called an $FEXPCOS$ model.

One of the nice features of $FEXP$ models is that Whittle's estimator of β can be obtained via generalized linear models (see Beran 1993), and, due to the linear form of $\log f$, the asymptotic covariance matrix Σ of $\sqrt{n}(\hat{\beta} - \beta)$ does not depend on β :

$$\Sigma = 4\pi D^{-1} \text{ with } D_{i,j} = \int_{-\pi}^{\pi} g_i(\lambda) g_j(\lambda) d\lambda \ (i, j = 0, 1, ..., p+1). \tag{5}$$

Since Y only depends on $w = \frac{1}{4}\Sigma_{22}$, it follows that the distribution of Y is nuisance parameter free.

3 Inference about μ for $FEXPCOS$ models

For an $FEXPCOS(p)$ model, D has zero elements everywhere except in the diagonal and in the second row and column. Also, $D_{11} = 2\pi$, $D_{22} = \pi^3/6$ and $D_{jj} = \pi$ ($j \geq 3$). For the remaining elements $D_{1j} = D_{j1}$ ($j \neq 0, 1$), we have

Lemma 1 *Consider an $FEXPCOS(p)$ model. Then, for $j \geq 3$,*

$$D_{2j}(p) = D_{j2}(p) = -\pi j^{-1}. \tag{6}$$

As noted in the previous section, the distribution of Y does not depend on β. However, the higher the value of p the higher the variance w of $\hat{d} = -\frac{1}{2}\hat{\beta}_1$, and thus the larger the confidence intervals for μ. The following proposition gives an explicit closed form formula for w and shows that w diverges to infinity linearily, as p increases:

Proposition 2 *Consider $FEXPCOS(p)$ models $(p = 0, 1, 2, ...)$ and the corresponding matrices $\Sigma(p)$ defined by (5). Let $a_j = j^{-1}$ $(j = 1, 2, ...)$, $a_o = 0$ and $w(p) = \frac{1}{4}\Sigma_{22}(p)$. Then*

$$w(p) = (\frac{\pi^2}{6} - \sum_{j=0}^{p} a_j^{-2})^{-1} \tag{7}$$

and

$$p \leq w(p) \leq p + 1 \tag{8}$$

Figure 1a displays $w(p)$ for $p = 0, 1, ..., 100$. A linear regression of $w(p)$ against p (see figure 1a) yields $R^2 = 1.00$ (rounded to two digits), a slope of one and an intercept of about 0.5. Thus, $w(p)$ is approximately in the middle between the two bounds p and $p + 1$. The plot of the residuals divided by $w(p)$ in Figure 1b shows that the linear approximation $w(p) \approx p + 0.5$ is almost exact for $p \geq 3$.

Since quantiles of Y depend on p only, they can be tabulated as a function of n and p. Table 1 gives the 95%-quantiles for $n = 1, 2, 3, ..., 20$ and $p = 0, 1, ..., 20$. Note that for high values of p the quantiles are very far from the corresponding standard normal quantiles, even for $n = 2000$.

4 Inference about μ for $FEXPO$ models

A further simplification can achieved by orthogonalizing $D(p)$ completely so that $w(p)$ does not depend on p anymore. The distribution of Y is then completely nuisance parameter free.

A first naive approach to orthogonalization is to start with an arbitrary $FEXP$ model (with functions $g_o, g_1, ..., g_{p+1}$) and then orthogonalize $g_o, g_1, ..., g_{p+1}$ sequentially by the Gram-Schmidt method. We would thus obtain an orthogonal basis of functions, say h_j $(j = 0, 1, ..., p + 1)$ and $\Sigma(p)$ would be diagonal. The question must be asked, however, whether every orthogonal basis of functions is statistically meaningful. The answer is no. For instance, Gram-Schmidt orthogonalization that starts with $h_o = 1$ and $h_1 = |1 - \exp(i\lambda)|$, leads to functions h_j $(j \geq 3)$ that diverge to plus or minus infinity at the origin. In the original definition of an $FEXP$

model, β_1 models long-range dependence whereas β_j $(j \geq 2)$ can be interpreted as short-memory parameters. This is no longer the case, if all functions (except h_o) are unbounded. Thus, Gram-Schmidt orthogolization destroys the statistically meaningful separation of short and long memory components in the parameter space. We therefore postulate that, in order to be statistically meaningful, a set of orthogonal functions h_j must be such that h_j $(j \geq 3)$ are bounded in $[-\pi, \pi]$. This can be achieved, for example, by the following

Algorithm 1 *Start with functions g_j such that $g_o = 1$ and $\int_{-\pi}^{\pi} g_o(\lambda) g_j(\lambda) d\lambda = 0$ $(j > 0)$. Define $a_j = \{\int g_1(\lambda) g_j(\lambda) d\lambda\}^{-1}$ $(j \geq 2)$, set $h_o = g_o$, $h_1 = g_1$ and carry out the following steps:*

- *Step 1: Define $u_j = a_j g_j - a_{j+1} g_{j+1}$ $(j \geq 2)$;*
- *Step 2: Apply Gram-Schmidt orthogonalization to u_j $(j \geq 2)$ to obtain orthogonal functions $h_2, ..., h_{p+1}$.*

Definition 1 *An $FEXP$ model with h_j defined by Algorithm 1 is called an orthogonal $FEXP$ model (or $FEXPO$ model).*

For $FEXPO$ models with $h_1 = |1 - \exp(i\lambda)|$, we have $w(p) = 6/\pi^2$ for all values of p. Note that this is equal to the smallest variance achievable by $FEXPCOS$ models. The quantiles of Y are the same as those for the $FEXPCOS(0)$ model (see table 1, $p = 0$).

5 Model choice for $FEXPO$ models

For $FEXPO(p)$ models, the distribution of Y does not depend on the chosen order p. However, for finite n, the value of the test statistic T and its finite sample distribution are influenced by p. Therefore, a suitable model choice criterion is needed. As an alternative to standard criteria (such as the AIC or BIC), the following simple model selection procedure can be used for $FEXPO$ models:

Algorithm 2 *1. Define a highest possible order P and a level of significance $0 < \alpha < 1$.*

2. *Estimate the parameter* $\beta^{(P)} = (\beta_o, \beta_1, ..., \beta_{P+1})^t$ *for the full model by Whittle's approximate maximum likelihood method.*

3. *Set* $\hat{\beta} = \hat{\beta}^{(P)}$ *and calculate, for* $j = 2, ..., P+1$, *the p-values* p_j *for testing* $H_o : \beta_j^{(P)} = 0$ *versus* $H_a : \beta_j^{(P)} \neq 0$. *If the* $p_j \leq \alpha P^{-1}$, *then set* $\hat{\beta}_j = 0$.

This procedure is justified by the fact that the components of $\hat{\beta}$ are asymptotically orthogonal to each other. The individual level of significance αP^{-1} corresponds to an exact Bonferroni correction. The probability of overfitting, i.e. keeping at least one unnecessary nonzero component, is equal to α. Note that the long-memory parameter is considered to be a "default" parameter here (otherwise testing with respect to β_2 would also have to be included).

Algorithm 2 has two advantages over other model choice criteria. It is *fast*, since estimation has to take place only once, and it is *non-hierarchical*, in the sense that a full comparison among all subset models is made. In contrast, most model choice criteria in time series analysis are applied in a hierarchical manner in that a comparison is made only among an increasing sequence of nested models. A full comparison of all possible subset models (as often done in regression) seems to be computationally infeasible when using the AIC or BIC.

6 Simulations

For $n = 400$ and $d = -0.3, 0$ and 0.3, the following Gaussian models were simulated:

1. Model 1: $FEXPCOS(0)$ model with $\beta = (1, -2d)$ (this is also an $FEXPO(0)$ model);

2. Model 2: $FEXPCOS(1)$ model with $\beta = (1, -2d, -0.5)$;

3. Model 3: $FEXPCOS(4)$ model with $\beta = (1, -2d, 0, 0, 0, -0.5)$;

4. Model 4: $FEXPO(1)$ model with $\beta = (1, -2d, 0.5)$;

5. Model 5: $FEXPO(4)$ model with $\beta = (1, -2d, 0, 0, 0, 0.5)$;

The number of simulations was 100. In all cases, $\mu = E(X_i)$ was equal to zero. The models with $p = 1$ and 4 were chosen such that for the same order p the spectral densities of the $FEXPCOS(p)$ and $FEXPO(p)$ model are qualitatively similar. To illustrate this, figures 2a through d display the spectral densities of Models 2 to 5 for the case with $d = 0.3$. For the $FEXPCOS$ models 1, 2 and 3, 95%-confidence intervals and tests (at a nominal level of 0.05) for μ were calculated for each series using a fitted $FEXPCOS(p)$ model, with p being estimated by the BIC (Schwarz 1978, Beran et al. 1999). The same was done for models 1, 4 and 5, using $FEXPO(p)$ models and algorithm 2 (with $\alpha = 0.05$).

Tables 2 and 3 give simulated rejection probabilities for testing $H_o : \mu = 0$ against $H_a : \mu \neq 0$. Table 4 gives summary statistics of the simulated lengths of 95%-confidence intervals for model 1. The results in tables 2 and 3 show that rejection probabilities are approximately correct. For $FEXPCOS$ models there seems to be a slight tendency to reject too often, in the case of long memory. This may be due to the fact that model selection plays a role for finite samples. This is less the case for $FEXPO$ models. For model 1, a direct comparison between the performance of $FEXPCOS$ and $FEXPO$ models is possible, since this process is included in both model classes: Table 4 indicates that in this case, $FEXPO$ models outperform $FEXPCOS$ models in the sense that confidence intervals tend to be shorter when based on an $FEXPO$ fit. In particular, for some simulated series very large $FEXPCOS$ confidence intervals occured, in contrast to $FEXPO$ intervals. Thus, orthogonalization tends to stabilize confidence intervals for μ.

7 Concluding remarks

In this note, a simple test procedure for inference about the location parameter of $FEXPCOS$ and $FEXPO$ models was discussed. Clearly there are many other ways of defining $FEXP$ models based on orthogonal functions. Which orthogonal functions are used may influence statistical inference, since each orthogonalization induces an intrinsic hierarchy of the resulting orthogonal functions h_j : the higher j the lower its "statistical priority". The same is true for models with nonorthogonal functions. The model choice criterion based on Bonferroni-corrected testing avoids this problem upto a certain degree, except that functions beyond a certain maximal order P are ignored. This criterion, however, is only reasonable for $FEXP$ models with orthogonal functions. Simulations where this criterion was applied to $FEXPCOS$ models showed a much higher variability in the estimates of β and of confidence intervals for μ.

Acknowledgement

This work was supported in part by a research grant of the University of Konstanz.

8 Appendix

Proof of Lemma 1: The result follows from Gradshteyn and Ryzhik (1979), formula 4.384.7

Proof of Proposition 1: Since $D_{1j} = 0$ $(j \neq 1)$, we may restrict attention to the submatrix $A = (a_{ij})_{i,j=1,\ldots,p+1}$ with $a_{ij} = d_{i+1,j+1}$. Now $A^{-1} = (\det A)^{-1} C^t$ where $C = (c_{ij})_{i,j=1,\ldots,p+2}$, $c_{ij} = (-1)^{i+j} \det A'_{ij}$ and A'_{ij} is the $p \times p$ submatrix of A obtained by cancelling the ith row and jth column. Note that $A_{11} = \pi^3/6$, $A_{jj} = \pi$ $(j \geq 2)$, $A_{1j} = A_{j1} = -\pi(j-1)^{-1}$ (Lemma 1) and A_{ij} is zero for all other indices. Then $\det A'_{11} = \pi^p$, $\det A'_{1j} = (-1)^{j+1}(j-1)^{-1}\pi^p$ $(j \geq 2)$, and $\det A = \sum_{j=1}^{p+1} (-1)^{1+j} a_{1j} \det A'_{1j} = \pi^{p+1}(\pi^2/6 - \sum_{j=1}^{p} j^{-2})$. Equation (7) then follows from $c_{11} = \pi^p$, $\Sigma_{22}(p) = 4\pi[A^{-1}]_{11}$ and $w(p) = \frac{1}{4}\Sigma_{22}(p)$. The lower and upper limits for w follow from $\sum_{j=1}^{\infty} j^{-2} = \pi^2/6$, and a Riemann approximation to $\sum_{j=p+1}^{\infty} j^{-2}$.

BIBLIOGRAPHY

Adenstedt, R.K. (1974) On large sample estimation for the mean of a stationary random sequence. Ann. Statist., **2**, 1095-1107.

Beran, J. (1989) A test of location for data with slowly varying serial correlations. Biometrika, **76**, 261-269.

Beran, J. (1993) Fitting long-memory models by generalized linear regression. Biometrika, **80, No. 4**, 817-822.

Beran, J. (1994) Statistics for long-memory processes. Chapman and Hall, New York.

Beran, J., Bhansali, R.J. and Ocker, D. (1998) On unified model selection for stationary and nonstationary short- and long-memory autoregressive processes. Biometrika, **Vol. 85, No. 4**, 921-934.

Bloomfield, P. (1973) An exponential model for the spectrum of a scalar time series. Biometrika, **60**, 217-226.

Diggle, P.J. (1990) Time series - a biostatistical introduction. Oxford: Oxford University Press.

Gradsteyn, I.S. and Ryzhik, I.M. (1979) Table of integrals, series, and products. London: Academic Press.

Granger, C.W.J., Joyeux, R. (1980) An introduction to long-range time series models and fractional differencing. J. Time Series Anal. **1** 15-30.

Hosking, J.R.M. (1981) Fractional differencing. Biometrika **68** 165-176.

Mandelbrot, B.B. and Wallis, J.R. (1969) Computer experiments with fractional Gaussian noises. Water Resources Research, **5, No. 1**, 228-267.

Samarov, A. and Taqqu, M.S. (1988) On the efficiency of the sample mean in long-memory noise. J. Time Series Analysis, **9**, 191-200.

Schwarz, G. (1978) Estimating the dimension of a model. Ann. Statist., **6**, 461-464.

Table 1: 95%, 97%, 99% and 99.5% quantiles for $Y = Z_1 n^{Z_2 \sqrt{w/n}}$ for $FEXPCOS(p)$ models.

95%−quantiles:						
	$n = 100$	200	400	1000	2000	∞
$p = 0$	1.84	1.77	1.73	1.68	1.67	1.645
1	2.13	1.97	1.85	1.76	1.71	1.645
2	2.44	2.17	1.98	1.82	1.76	1.645
3	2.76	2.37	2.11	1.90	1.80	1.645
4	3.11	2.59	2.24	1.96	1.84	1.645
5	3.45	2.81	2.37	2.03	1.88	1.645
6	3.81	3.03	2.51	2.10	1.92	1.645
7	4.18	3.26	2.65	2.17	1.96	1.645
8	4.56	3.50	2.79	2.24	2.01	1.645
9	4.97	3.73	2.93	2.31	2.05	1.645
10	5.41	3.97	3.08	2.39	2.09	1.645
11	5.85	4.21	3.23	2.46	2.13	1.645
12	6.31	4.47	3.38	2.53	2.18	1.645
13	6.79	4.73	3.53	2.60	2.22	1.645
14	7.29	5.01	3.67	2.68	2.26	1.645
15	7.81	5.30	3.83	2.75	2.30	1.645
16	8.34	5.60	3.98	2.83	2.35	1.645
17	8.93	5.88	4.14	2.90	2.39	1.645
18	9.52	6.19	4.30	2.98	2.44	1.645
19	10.15	6.50	4.46	3.06	2.48	1.645
20	10.76	6.83	4.62	3.14	2.52	1.645

(Table 1 continued)

97.5%−quantiles:						
$p = 0$	2.32	2.20	2.11	2.04	2.01	1.96
1	2.87	2.58	2.35	2.17	2.09	1.96
2	3.45	2.94	2.61	2.30	2.17	1.96
3	4.08	3.33	2.84	2.44	2.25	1.96
4	4.74	3.72	3.08	2.57	2.33	1.96
5	5.43	4.16	3.33	2.70	2.41	1.96
6	6.18	4.60	3.58	2.82	2.49	1.96
7	7.01	5.05	3.85	2.95	2.58	1.96
8	7.88	5.52	4.12	3.08	2.65	1.96
9	8.78	6.00	4.41	3.22	2.73	1.96
10	9.73	6.55	4.69	3.35	2.80	1.96
11	10.81	7.10	4.97	3.48	2.88	1.96
12	11.96	7.68	5.28	3.62	2.96	1.96
13	13.19	8.25	5.57	3.75	3.04	1.96
14	14.52	8.86	5.88	3.90	3.11	1.96
15	15.87	9.53	6.22	4.05	3.20	1.96
16	17.30	10.18	6.57	4.20	3.29	1.96
17	18.76	10.91	6.93	4.35	3.36	1.96
18	20.37	11.67	7.29	4.50	3.44	1.96
19	22.11	12.45	7.66	4.65	3.52	1.96
20	23.85	13.26	8.03	4.80	3.61	1.96

(Table 1 continued)

99%−quantiles:						
	$n = 100$	200	400	1000	2000	∞
$p = 0$	2.96	2.76	2.59	2.46	2.39	2.33
1	3.94	3.41	3.01	2.70	2.55	2.33
2	5.04	4.09	3.45	2.93	2.70	2.33
3	6.25	4.81	3.88	3.15	2.83	2.33
4	7.54	5.60	4.34	3.40	2.97	2.33
5	9.03	6.41	4.81	3.63	3.11	2.33
6	10.66	7.26	5.31	3.86	3.26	2.33
7	12.47	8.19	5.82	4.11	3.40	2.33
8	14.54	9.21	6.34	4.34	3.54	2.33
9	16.77	10.30	6.87	4.59	3.68	2.33
10	18.99	11.41	7.46	4.84	3.83	2.33
11	21.60	12.70	8.04	5.11	3.97	2.33
12	24.55	14.04	8.67	5.38	4.12	2.33
13	27.60	15.47	9.33	5.65	4.27	2.33
14	30.81	16.98	10.04	5.93	4.42	2.33
15	34.50	18.39	10.74	6.21	4.56	2.33
16	38.62	20.04	11.47	6.48	4.72	2.33
17	43.07	21.81	12.26	6.76	4.86	2.33
18	47.71	23.74	13.15	7.05	5.02	2.33
19	52.68	25.79	13.99	7.37	5.19	2.33
20	58.19	27.82	14.90	7.67	5.35	2.33

(Table 1 continued)

99.5%−quantiles:						
$p = 0$	3.49	3.18	2.99	2.80	2.71	2.58
1	4.92	4.10	3.56	3.10	2.92	2.58
2	6.53	5.13	4.16	3.44	3.10	2.58
3	8.37	6.16	4.83	3.76	3.29	2.58
4	10.53	7.35	5.50	4.08	3.51	2.58
5	12.89	8.61	6.17	4.41	3.69	2.58
6	15.62	10.07	6.93	4.79	3.90	2.58
7	18.74	11.57	7.70	5.16	4.09	2.58
8	22.07	13.22	8.52	5.51	4.28	2.58
9	25.94	15.03	9.40	5.85	4.50	2.58
10	30.36	16.98	10.40	6.22	4.72	2.58
11	35.25	19.11	11.35	6.63	4.95	2.58
12	40.28	21.36	12.37	7.05	5.18	2.58
13	46.30	23.76	13.42	7.44	5.39	2.58
14	52.99	26.32	14.61	7.86	5.62	2.58
15	59.83	29.13	15.76	8.29	5.80	2.58
16	67.49	32.27	17.10	8.75	6.02	2.58
17	75.32	35.65	18.41	9.22	6.26	2.58
18	84.57	38.97	19.74	9.70	6.50	2.58
19	95.36	42.62	21.29	10.28	6.76	2.58
20	106.55	46.69	22.67	10.74	7.00	2.58

Table 2: Rejection probabilities of the test for the location parameter based on $FEXPCOS$ models. The results are based on 100 simulations of an $FEXPCOS$ model with $\beta = (1, -2d)$ (Model 1), $\beta = (1, -2d, -0.5)$ (Model 2) and $\beta = (1, -2d, 0, 0, 0, -0.5)$ (Model 3) respectively.

	$d = -0.3$	$d = 0$	$d = 0.3$
Model 1	0.06	0.06	0.03
Model 2	0.01	0.07	0.12
Model 3	0.00	0.03	0.11

Table 3: Rejection probabilities of the test for the location parameter based on $FEXPO$ models. The results are based on 100 simulations of an $FEXPO$ model with $\beta = (1, -2d)$ (Model 1), $\beta = (1, -2d, 0.5)$ (Model 4) and $\beta = (1, -2d, 0, 0, 0, 0.5)$ (Model 5) respectively.

	$d = -0.3$	$d = 0$	$d = 0.3$
Model 1	0.05	0.06	0.03
Model 2	0.02	0.08	0.08
Model 3	0.02	0.02	0.05

Table 4: Simulated length of confidence intervals for an $FEXPCOS(0)$ model using $FEXPCOS$ and $FEXPO$ fits respectively. Notation: Q_1 =lower quartile, Q_2 =upper quartile, M =median.

	minimum	maximum	Q_1	Q_2	mean	M	std. dev.
$d = -0.3$							
$FEXPCOS$	0.15	0.84	0.18	0.24	0.22	0.20	0.09
$FEXPO$	0.13	0.30	0.17	0.23	0.20	0.19	0.04
$d = 0$							
$FEXPCOS$	0.37	2.25	0.71	1.00	0.90	0.84	0.32
$FEXPO$	0.32	1.51	0.71	0.98	0.85	0.84	0.23
$d = 0.3$							
$FEXPCOS$	2.76	106.77	4.51	7.73	7.51	5.87	10.49
$FEXPO$	2.60	14.22	4.47	7.18	6.07	5.88	2.15

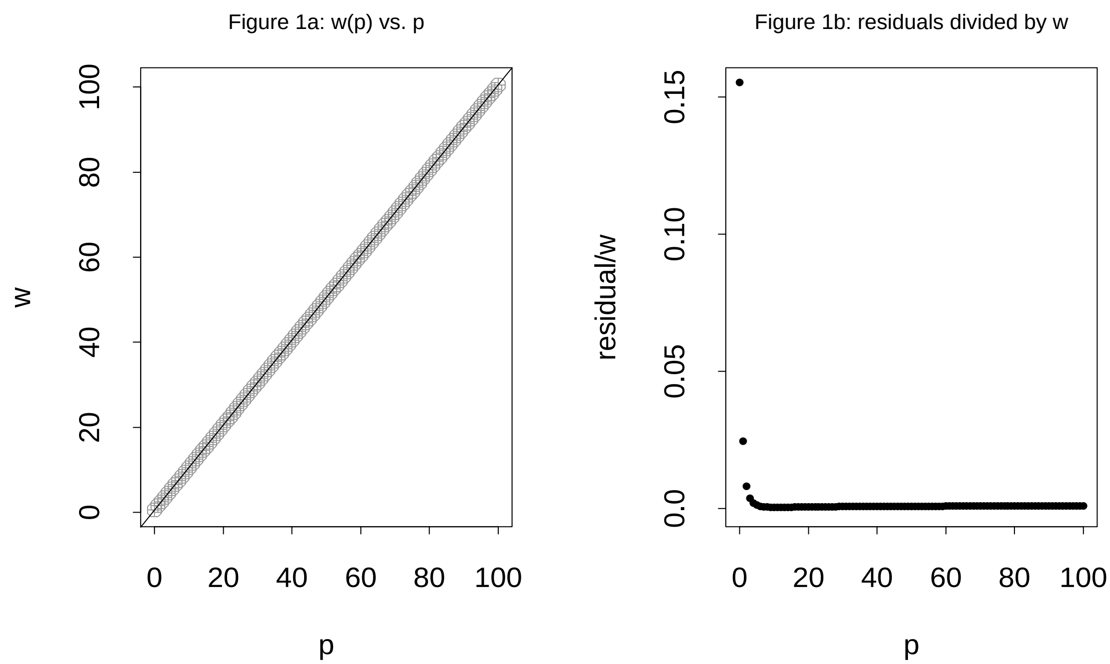

Figure 1: Asymptotic variance $w(p)$ of $\sqrt{n}(\hat{d} - d)$ as a function of p for $FEXPCOS(p)$ models (Figure 1a). Figure 1b shows the residuals of the least squares fit divided by $w(p)$, plotted against p.

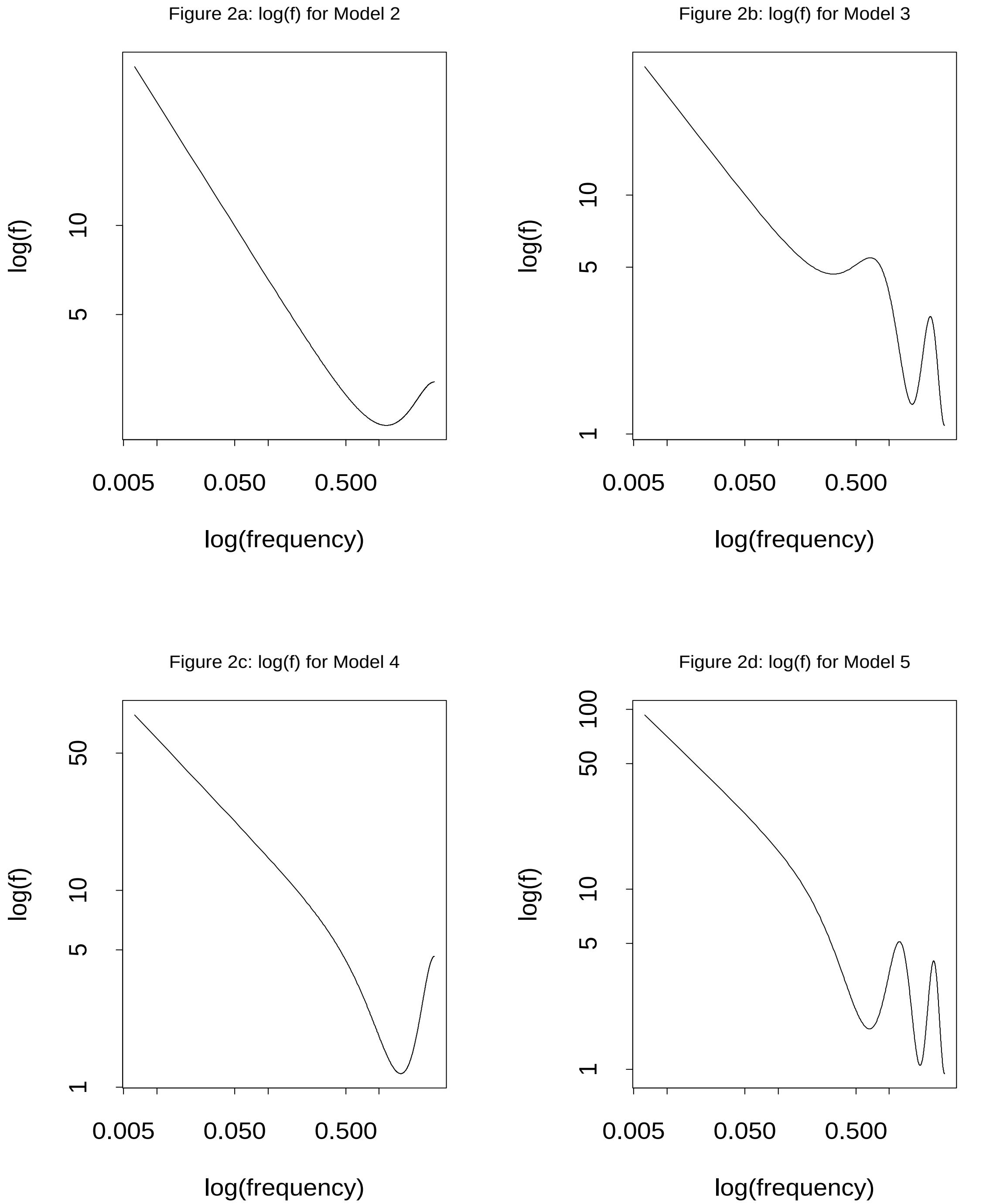

Figure 2: Spectral densities (in log-log-coordinates) of Models 2 (Figure 2a), 3 (Figure 2b), 4 (Figure 2c) and 5 (Figure 2d) used in the simulations, for the case where $d = 0.3$.