Project 1 - SPATIAL STATISTICS

Kwaku Peprah Adjei, Amir Ahmed

Problem 1

Let $\{r(x) : x \in \mathrm{D} : [1, 50] \subset \mathbb{R}^1\}$. Assume it is modelled as a stationary one dimensional Gaussian random field with the following parameters:

$$\begin{aligned} E\{r(x)\} &= \mu_r = 0 \\ Var\{r(x)\} &= \sigma_r^2 \\ Corr\{r(x), r(x')\} &= \rho_r(\tau) \\ \tau &= \frac{|x - x'|}{10} \end{aligned} \tag{1}$$

where $\rho_r(\tau)$ is the spatial correlation function. Discretize D as $L = \{1, 2, \ldots, 50\}$, we will later look at the discretized Gaussian random field $\{r(x) : x \in L\}$.

Problem 1a

To ensure the positive definetness of all covariance matrices, we rely on having a positve definite corrolation function. A positive definite correlation function guarantees that the covariance matrix of our Gaussian random field is well defined for all choices of observation points and for all dimensions.

A function $c(\boldsymbol{\tau}) : \mathbb{R}^q$ i positive definite if the following is satisfied:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j c(\mathbf{x}_i - \mathbf{x}_j) \geq 0, \quad \begin{array}{l} \forall \alpha_1, \ldots, \alpha_n \in \mathbb{R} \\ \forall n \in \mathbb{N}, \quad n \geq 2 \\ \forall \mathbf{x}_1, \ldots \mathbf{x}_n \in \mathbb{R}^q \end{array} \tag{2}$$

Any $n \times n$ matrix, $\mathbf{Q} = [c(\mathbf{x}_i - \mathbf{x}_j)]_{ij}$ constructed with a positive definite $c(\boldsymbol{\tau})$ with vectors $\{\mathbf{x}_i\}_{i=1}^n$ chosen as in (2) would clearly be positive definite. As for all $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$ we have:

$$\boldsymbol{\alpha}^T \mathbf{Q} \boldsymbol{\alpha} = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j c(\mathbf{x}_i - \mathbf{x}_j) \geq 0 \tag{3}$$

So it would be safe to use $\mathbf{Q}$ as a covariance matrix, and it would satisfy the requirements of the multivariate normal distribution.

We now go on to look at the the powered exponential correlation function:

$$\rho(\tau) = \exp(-\tau^\nu), \quad \nu \in (0, 2] \tag{4}$$

for later case studies we will look at parameters $\nu_r \in \{1, 1.9\}$. We note that and increased ν ($\nu >> 1$) would lower the exponent and thus give less correlation.

We also want to study the Matern correlation function

$$\rho_r(\tau) = \frac{2^{1-\nu}}{\Gamma(\nu)} \tau^v \mathcal{B}_\nu(\tau), \quad \nu \in \mathbb{R}_+ \tag{5}$$

$\mathcal{B}_\nu$ is the bessel function. We want to look at the matern with parameters $\nu_r \in \{1, 3\}$.

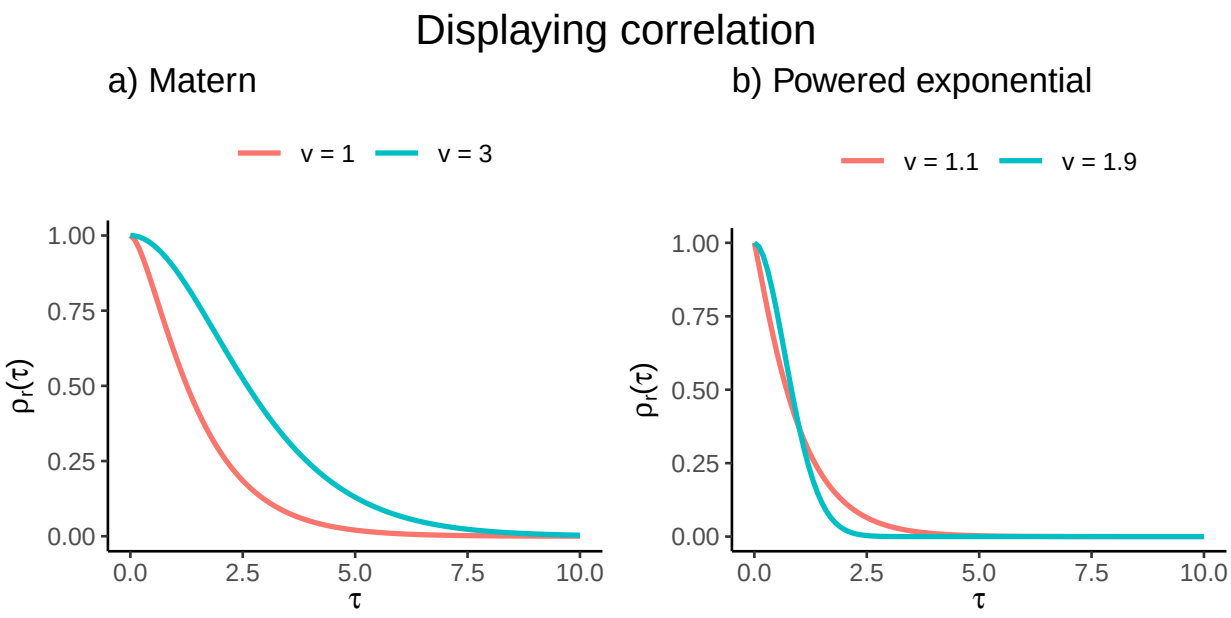


Figure 1: Plotting matern and exponential correlation

For both correlation functions use $\sigma_r^2 \in \{1, 5\}$

The two correlation functions are plotted in Figure 1 for $\tau \in \mathbb{R}_{\oplus}$. We note that these functions both are positive definite. From the figures we also note that both correlation functions for the different parameters seems to satisfy ergodicity, correlation drops to zero the further apart to points are, this is a neccesary trait in our Gaussian RF.

For the matern corrolation we see that an increase of ν lead to slower fall in correlation with distance. We also note that the Matern seems to drop slower in correlation than the exponential correlation function for the given parameters.

For higher values of ν and when $\tau >> 1$ the powered exponential seems to drop faster in correlation than for higher ν. The oppsite seems to be the trend when $\tau << 1$

The variogram function is defined as:

$$\begin{aligned} \gamma_r(\tau) &= \frac{1}{2} Var\{r(x) - r(x')\} \\ &= \frac{1}{2}(Var\{r(x)\} + Var\{r(x')\} - 2Corr\{r(x), r(x'))\} \\ &= \sigma_r^2(1 - \rho_r(\tau)) \end{aligned} \tag{6}$$

Looking at (6) we see that if $\sigma_r^2 = 1$ then the corrolation function and the variogram functions would respectively increase and decrease at same rate in with respect to an increase in τ. In essence the variogram function tells us how much variance we have in our estimation of x' when it is τ away from a observed point x

We display the variograms for our model parameters in Figure 2.

In both figures we see that the value of σ^2 sets a roof on the variance of points far away. Higher values of σ^2 also seems to increase the max achieved variance. For the matern variogram increased ν seems to increase corrolation to neighbouring points.

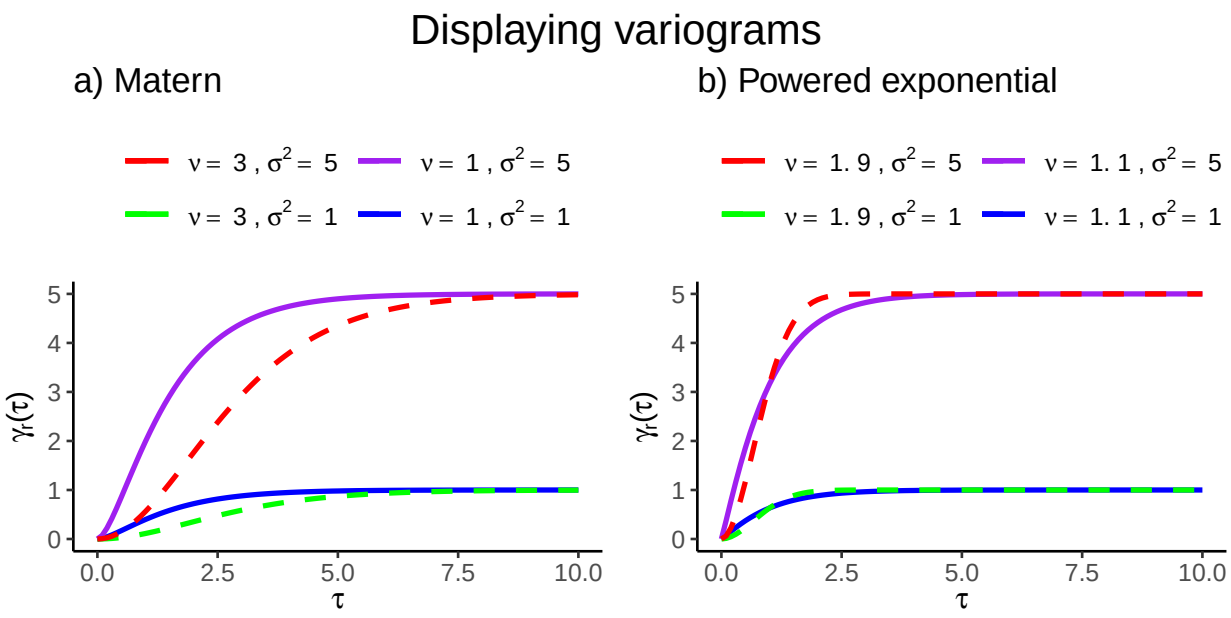


Figure 2: Plotting variograms

Problem 1b

We use the corrolation function $\rho_r(\tau)$ to construct a corrolation matrix. The corrolation matrix would have the following form:

$$\mathbf{\Sigma}_r^\rho = \begin{bmatrix} \rho_r(1,1) & \rho_r(1,2) & \dots & \rho_r(1,50) \\ \rho_r(2,1) & \rho_r(2,2) & \dots & \rho_r(2,50) \\ \vdots & \vdots & \ddots & \vdots \\ \rho_r(50,1) & \rho_r(50,2) & \dots & \rho_r(50,50) \end{bmatrix} \tag{7}$$

$\rho_r(x, x')$ denotes $\rho_r(\tau) = \rho_r(|x - x'|/10)$. With variance σ_r^2 the prior get the covariance matrix $\mathbf{\Sigma}_r = \sigma_r^2 \mathbf{\Sigma}_r^\rho$. From (1) we have an expected value of:

$$\boldsymbol{\mu}_r = (0, \dots, 0)^T \tag{8}$$

Which gives the prior a distribution of:

$$\phi_{50}(\mathbf{r}; \boldsymbol{\mu}_r, \mathbf{\Sigma}_r) \tag{9}$$

Which has pdf:

$$(2\pi)^{-\frac{50}{2}} \det(\mathbf{\Sigma}_r)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\mathbf{x}^T \mathbf{\Sigma}_r^{-1} \mathbf{x}\right) \tag{10}$$

We simulate four realisations of the field for the different parameters.

The results are plotted in Figure 3 and Figure 4. We note that an increased variance seems to increase fluctuation (comparing c) d) g) and h) to the others). For the matern we also see that an increased ν seems to make the process smoother. For both correlation functions increased ν seems to increase "air-time" when the process is strays away from expected 0.

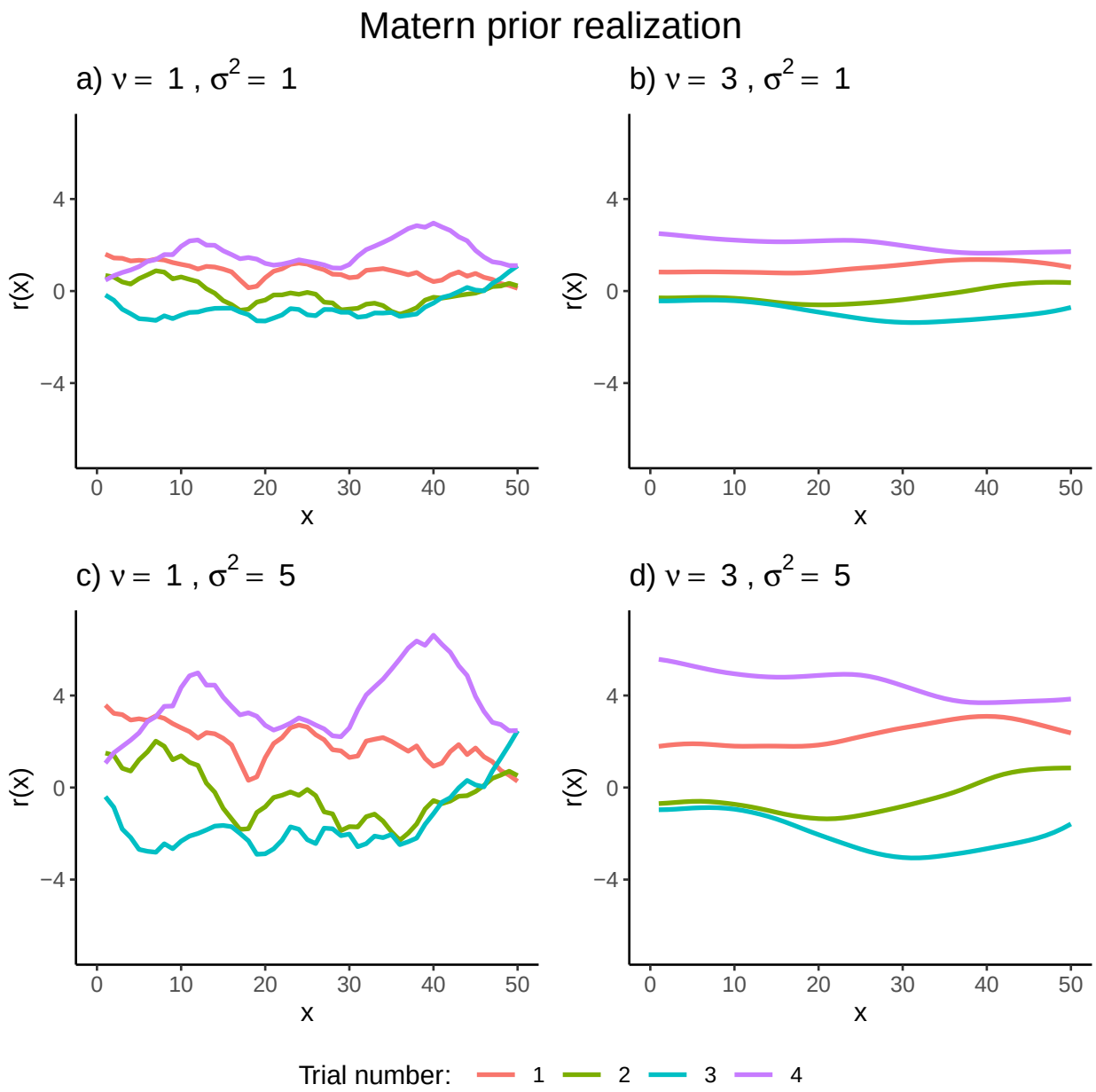


Figure 3: Realizisations of prior, matern

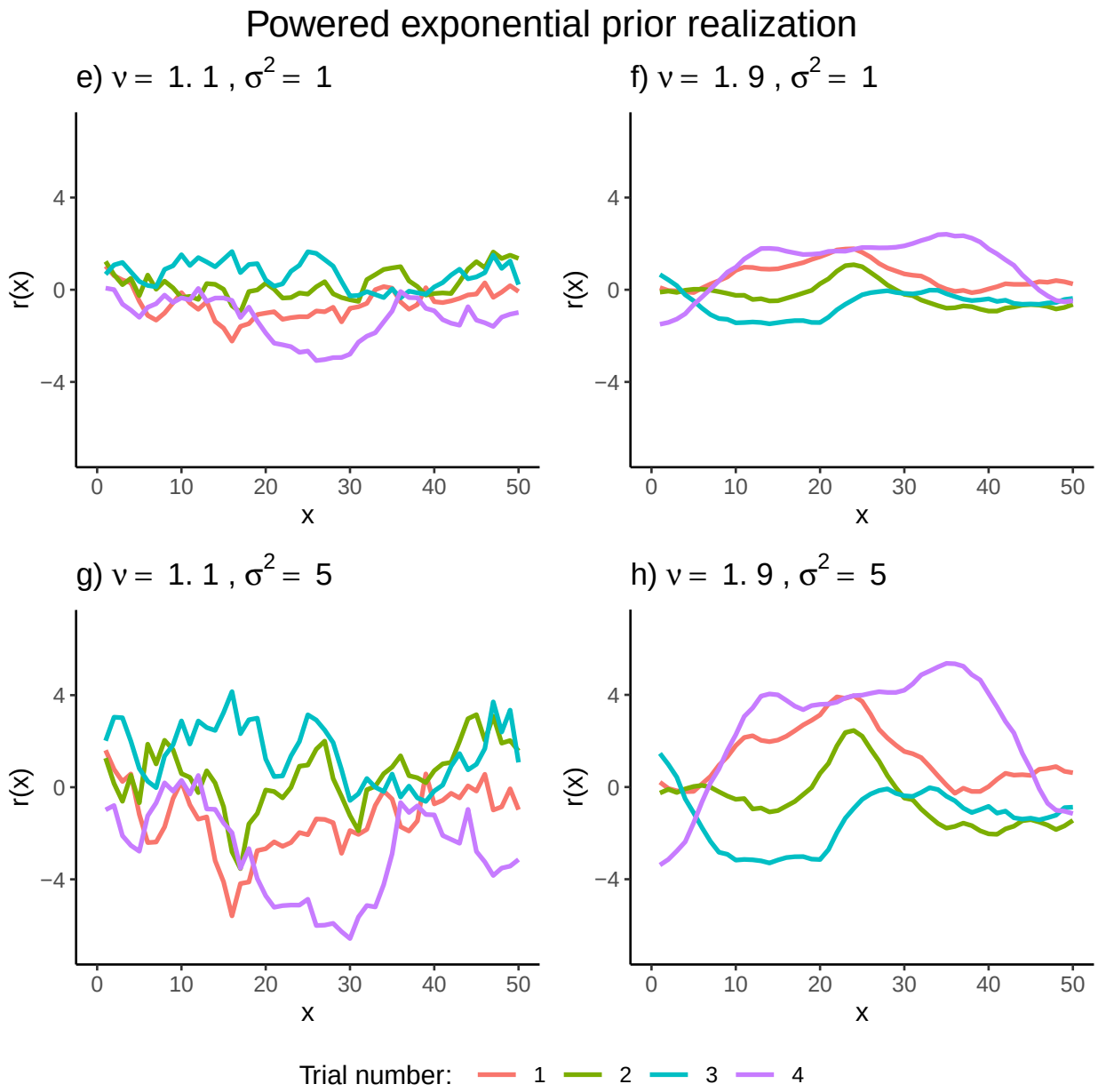


Figure 4: Realizisations of prior, powered exponential

Problem 1c

We now want to develop a posterior model. Assume that we have observed the values at $x \in \{10, 25, 30\}$. Organise them as:

$$\{d(x); x \in \{10, 25, 30\} \subset L\} \tag{11}$$

We also assume we have an observation error $\sigma_\epsilon^2 = \{0, 0.25\}$. In general we have

$$d(x) = r(x) + \epsilon(x), \quad x \in \{10, 25, 30\} \tag{12}$$
$$\epsilon(x) \sim N(0, \sigma_\epsilon^2), \text{iid} \tag{13}$$
$$r(x) \text{ and } \epsilon(x) \text{ independent} \tag{14}$$
$$\epsilon(x) \text{ and } \epsilon(x') \text{ independent identically distributed} \tag{15}$$

As we both $\epsilon(x)$ and $r(x)$ are Gaussian, a linear product of the two would also be.

We further note:

$$E(d(x)) = E(r(x)) + E((\epsilon(x))) = 0 \tag{16}$$

When $x \neq x'$

$$\begin{aligned} \text{Cov}(d(x), d(x')) &= \text{Cov}(r(x) + \epsilon(x), r(x') + \epsilon(x')) \\ &= \text{Cov}(r(x), r(x')) + \text{Cov}(\epsilon(x), \epsilon(x')) \\ &= \sigma_r^2 \rho_r(\tau) \end{aligned}$$

And:

$$\begin{aligned} \text{Cov}(d(x), d(x)) &= \text{Cov}(r(x) + \epsilon(x), r(x) + \epsilon(x)) \\ &= \text{Cov}(r(x), r(x)) + \text{Cov}(\epsilon(x), \epsilon(x)) \\ &= \sigma_r^2 \rho_r(\tau) + \sigma_e^2 \end{aligned} \tag{17}$$

We organise the observed points in a $k \times 1$ vector,$\mathbf{y} = (10, 25, 30)^T$, where $k = 3$. Further let the observed values of the random field be:

$$\mathbf{d}(\mathbf{y}) = (d(10), d(25), d(30))^T \tag{18}$$

If we denote $\mathbf{\Sigma}_\rho^d$ as the correlation matrix generated by $\rho(\tau)$ between the points $\mathbf{x}$ we see:

$$\text{Cov}(\mathbf{d}(\mathbf{x}), \mathbf{d}(\mathbf{x}')) = \sigma_e^2 \mathbf{I}_k + \sigma_r^2 \mathbf{\Sigma}_d^\rho \tag{19}$$

Using the calculations above $\mathbf{d}$ thus have the following pdf:

$$p(\mathbf{d}(\mathbf{x}) | \sigma_r^2, \sigma_e^2) = (2\pi)^{k/2} \det(\mathbf{\Sigma}_\rho^d)^{-1/2} \exp(\frac{1}{2} \mathbf{d}^T (\mathbf{\Sigma}_\rho^d)^{-1} \mathbf{d}) \tag{20}$$

The value σ_e^2 would then have the following likelihood function:

$$L(\sigma_e^2|\mathbf{d}(\mathbf{x}), \sigma_r^2) = p(\mathbf{d}(\mathbf{x})|\sigma_r^2, \sigma_e^2) \tag{21}$$

The integral

$$\int_0^\infty L(\sigma_e^2|\mathbf{d}(\mathbf{x}), \sigma_r^2)d\sigma^2 = \int_0^\infty p(\mathbf{d}(\mathbf{x})|\sigma_r^2, \sigma_e^2)d\sigma^2 \tag{22}$$

goes to infinity, thus $L(\sigma_e^2|\mathbf{d}(\mathbf{x}), \sigma_r^2)$ is not a pdf. If we try to estimate σ_ϵ^2 by the expected value of σ_e^2 derived when believing $L(\sigma_\epsilon^2|\cdot)$ is a pdf, we would expect values that goes to infinity, which would be wrong.

Problem 1d

Assume now that $\sigma_r^2 = 5$ and that we have observed $\mathbf{d}(\mathbf{y})$ with error $\sigma_\epsilon^2 =\in \{0, 0.25\}$. Want to find the distribution of $[\mathbf{r}|\mathbf{d}]$. Know that both $\mathbf{r}$ and $\mathbf{d}$ are multivariate normal. Thus conditioning $\mathbf{r}$ on $\mathbf{d}$ would give a multivariate normal distribution with the following parameters:

$$\boldsymbol{\mu}_{\mathbf{l}|\mathbf{d}} = \boldsymbol{\mu}_r + \boldsymbol{\Sigma}_{\mathbf{r},\mathbf{d}}\boldsymbol{\Sigma}_{\mathbf{d}}^{-1}(\mathbf{d} - \boldsymbol{\mu}_{\mathbf{d}}) \tag{23}$$

$$\boldsymbol{\Sigma}_{\mathbf{r}|\mathbf{d}} = \boldsymbol{\Sigma}_{\mathbf{r}} - \boldsymbol{\Sigma}_{\mathbf{r},\mathbf{d}}\boldsymbol{\Sigma}_{\mathbf{d}}^{-1}\Sigma_{\mathbf{d},\mathbf{r}} \tag{24}$$

Note $\Sigma_{\mathbf{r},\mathbf{d}} = \Sigma_{\mathbf{d},\mathbf{r}}^T$ and they are respectively 50×3 and 3×50 where the first has elements $[\text{Cov}\{r(x_i), d(x_j)\}]_{ij}$ where:

$$\begin{aligned} \text{Cov}\{r(x_i), d(x_j)\} &= \text{Cov}\{r(x_i), r(x_j) + \epsilon(x_j)\} \\ &= \text{Cov}\{r(x_i), r(x_j)\} \\ &= \sigma_r^2\rho_r(x_i, x_j) \end{aligned} \tag{25}$$

Assuming we have realisations of $\mathbf{d}$ from simulated data we can study how $\mathbf{r}|\mathbf{d}$ acts.

We compute two corresponding predictions for the spatial variable $\{\hat{r}(x); x \in L\}$ with associated 0.9 predictions intervals and display them in Figure 5. We use the same realisations Trial 1 in d) in Figure 3.

Here we have parameters $\nu_r = 3$ and $\sigma_r^2 = 5$. As discussed earlier, we would with these values see much corrolation. From the variogram, points with distance up to about 20 in distance would impact each other, this is reflected in the figure, as the observation points shifts large parts of the process away from 0. The high correlation decreases the width of the 90% confidence intervals close to the observations. In the interval $x \in (25, 30)$ the 90% band is very narrow in the exact observation case, comapred to the one with observation error. Comparing no variance in observation to variance the points close to the observation points in no observation error is of course more accurate, but further away i.e. at point 20 the increased variance seem to have little impact on the confidence band. The real process behind, keeps within 90% although we would expect more of them to stray of, as we have 47 non observed-points.

Problem 1e

We now simulate 100 realisation's with the same parameters as in 1d) and calculate empirical mean and empirical 90% confidence bands.

The results are illustrated in Figure 6. We also include the theoretical plots that the simulations are drawn from. For both with observation error and without observations error the empirical estimations seems to do well, and overlaps quite well with the theoretical bands. The largest differences between the two are seen furthest away from the observations points. Between $x \in (10, 25)$ theoretical bands seems to have a slight

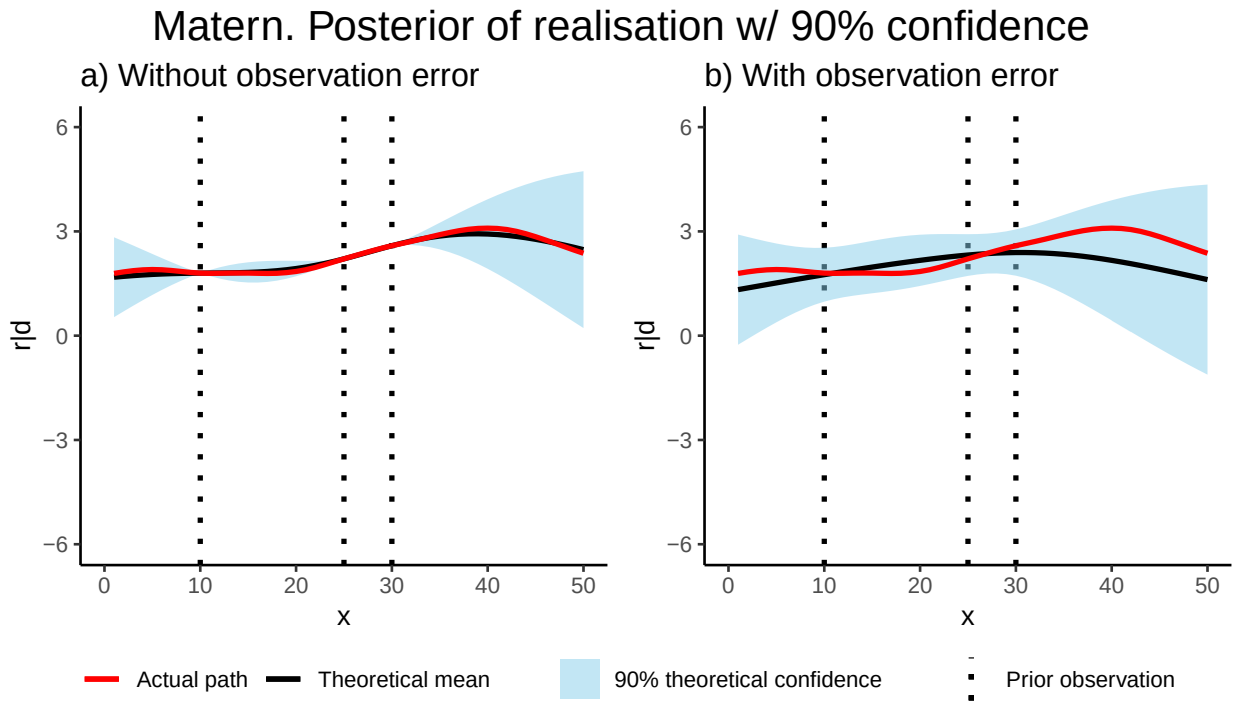


Figure 5: Posterior estimation, matern $\sigma^2 = 5$, $\nu = 3$

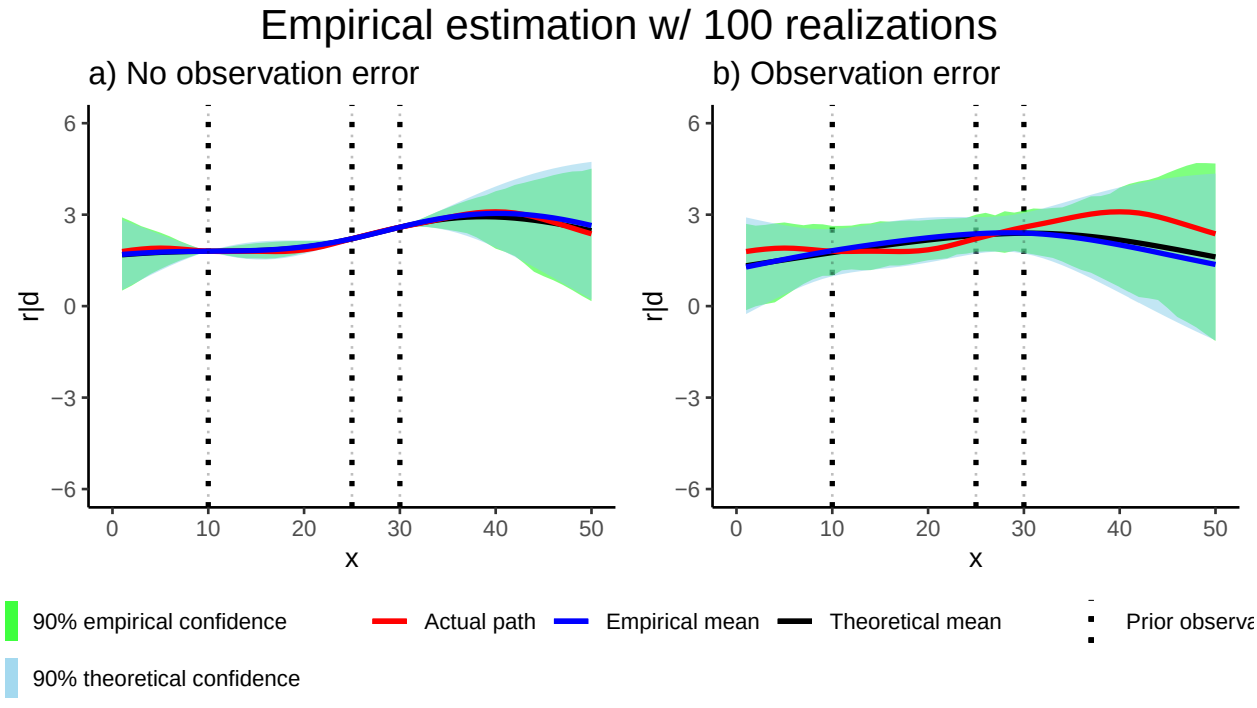


Figure 6: Posterior estimation 100 realisation, matern $\sigma^2 = 5$, $\nu = 3$

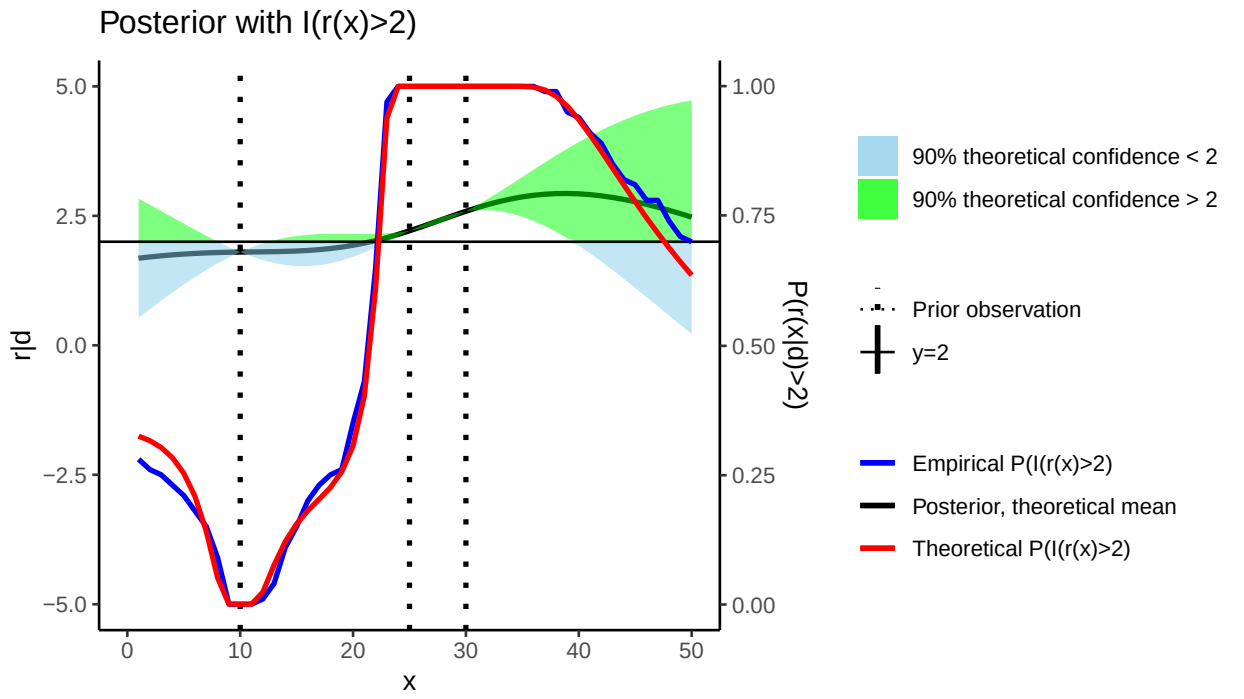


Figure 7: Posterior estimation 100 realisation, matern $\sigma^2 = 5$, $\nu = 3$

tendency to be wider that the empirical bands, however the trend is not too clear. Close to the observation points the bands seems to be equally wide without observation error, with obserbation error there seems to be slightly more variation.

Problem 1f

Want to evaluate

$$A_r = \int_D I(r(x) > 2)dx \tag{26}$$

The situation is illustrated in Figure 7

A prediction of $\hat{A}$ using our 100 realizations from the posterior model with σ_ϵ^2 is:

$$\hat{A}_r = \frac{1}{100} \sum_{i=1}^{100} \sum_{x \in L} I(r_i(x) > 2) \tag{27}$$

Applying this to our realisation we get: $\hat{A} = 30.16$ and $\text{Var}(\hat{A}) = 32.519596$

Where we use the approximation:

$$A_r = \int_D I(r(x) > 2)dx \approx \sum_{x \in L} I(r(x) > 2)dx \tag{28}$$

Another estimator is:

$$\tilde{A}_r = \sum_{x \in L} I(\hat{r}(x) > 2) \tag{29}$$

With our data we get: $\tilde{\text{A}} = 29$

We note that in our case $\hat{\text{A}}_r > \tilde{\text{A}}_r$.

If we let:

$$g(r(x)) = \int_D I(r(x) > 2) dx \tag{30}$$

Then $g(x)$ is convex as it is an integral of a positive value.

Using Jensen's Inequality we thus get:

$$Eg(r(x)) \geq g(Er(x)) = \int_D I(Er(x) > 2) = \int_D I(\hat{r}(x) > 2) \tag{31}$$

The same would be the case for the discretized stimators, i.e. we have:

$$E\hat{A}_r \geq \tilde{A}_r$$

And is why we expect $\hat{A}_r > \tilde{A}_r$. Which is also what we observed.

Another way to look at the problem is the following:

Let Y_{ij} be that event that the j-th observation at the i-th measurement point ($0 \leq i \leq 50$) is over or under 2 in the posterior model. We define:

$$Y_{ij} := I(r(X_{ij}) > 2) \tag{32}$$

Where X_{ij} is the j-th observation positionat the i-th realisation of the posterior. As the 100 posterior realisations are identically distributed and independent. Each Y_{ij} is independent grouped for position j. Then Y_{ij} is Bernoulli distributed. With probability p_i of being higher than 2. Where:

$$p_i = 1 - \Phi\left(\frac{2 - \mu_i}{\sigma_i}\right) \tag{33}$$

Where $\mu_i = \left[\boldsymbol{\mu}_{l|d}\right]_i$ and $\sigma_i^2 = \left[\boldsymbol{\Sigma}_{l|d}\right]_{ii}$

We know that $\hat{p}_i = 1/100 \sum_{i=1}^{100} Y_{ij}$ is an unbiased estimator for p_i

The theoretical and estimated values of p_i are illustrated in Figure **??**, and seem to match the situation and each other quite well. They are scaled such that the height of the plot (-5, 5) would be $p_i = 1$

This thinking can also be applied to estimating A_r

One alternative $\hat{A}_r$ is:

$$\hat{A}_r = \frac{1}{100} \sum_{i}^{50} w \sum_{j}^{100} I(X_{ij} > 2) \tag{34}$$

We note that:

$$\begin{aligned} E\hat{A}_r &= \frac{1}{100}\sum_{i}^{50} w \sum_{j}^{100} EI(X_{ij} > 2) \\ &= \frac{1}{100}\sum_{i=1}^{50} w \sum_{j=1}^{100} \left[1 - \Phi\left(\frac{2-\mu_i}{\sigma_i}\right)\right] \\ &= \sum_{i}^{50} w p_i \end{aligned} \tag{35}$$

Whre w is the width between each observation, in our case $w = 1$. We see that when number of observations goes to infinity, the above sum goes toward the integral we want to estimate.

This can also be used to show that we expect $\hat{A} > \tilde{A}$. By using that:

$$P(\hat{r}(x_i) > 2) = 1 - \Phi\left(\frac{2-\mu_i}{\sigma_i}\sqrt{n}\right) < 1 - \Phi\left(\frac{2-\mu_i}{\sigma_i}\right) \tag{36}$$

The inequality can be easily proven by using this.

Problem 1g)

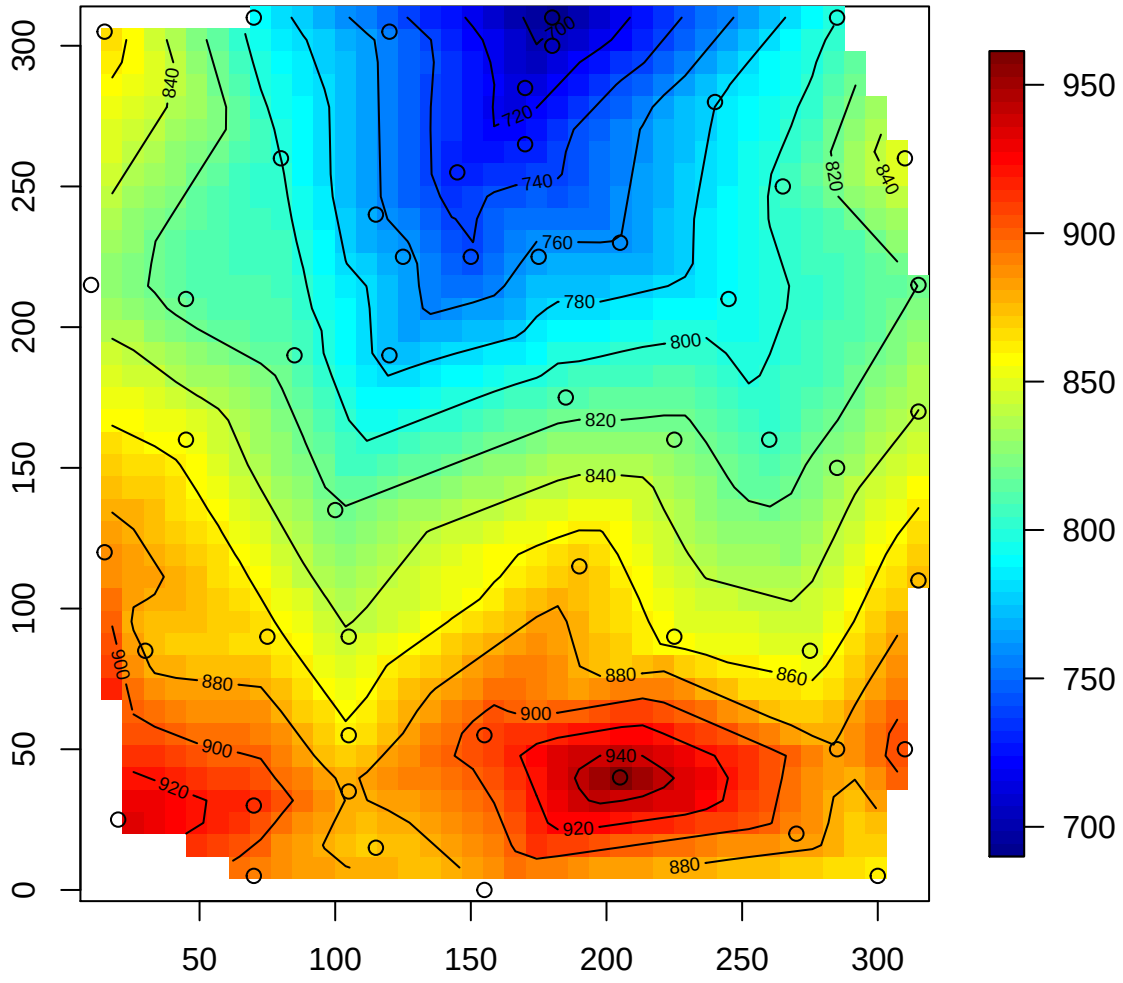
We have studied the matern and the powered exponential correlation function, and applied them to simulated data, to estiamte posteriors and their confidence. We find it interesteing how a change in the ν seem to change the differentiability of the end smoothness. Overall the empirical estimation methods seems to work well and are consistent with theory.

Problem 2 Gaussian RF - real data

Given the domain $D = [(0, 315) \times (0, 315)] \subset \mathbb{R}^{\not\in}$. We let $\mathbf{d} = r(\mathbf{x_1^0}), ..., r(\mathbf{x_{52}^0}))^T$.

Problem 2a: Display of the data

The data is displayed with an image plot, a contour plot and the exact points as shown in the figure **??** below.



It was observed that the data points did not move in the same direction as with the x and y cordinates (see figure 8 a;b). This suggest that the data is not mean stationary. Moreover, a density plot for the data points in figure 8 c) shows a skewness in the data, making the Guassianity assumption doubtful. Hence a stationary Gaussian RF may not be appropriate.

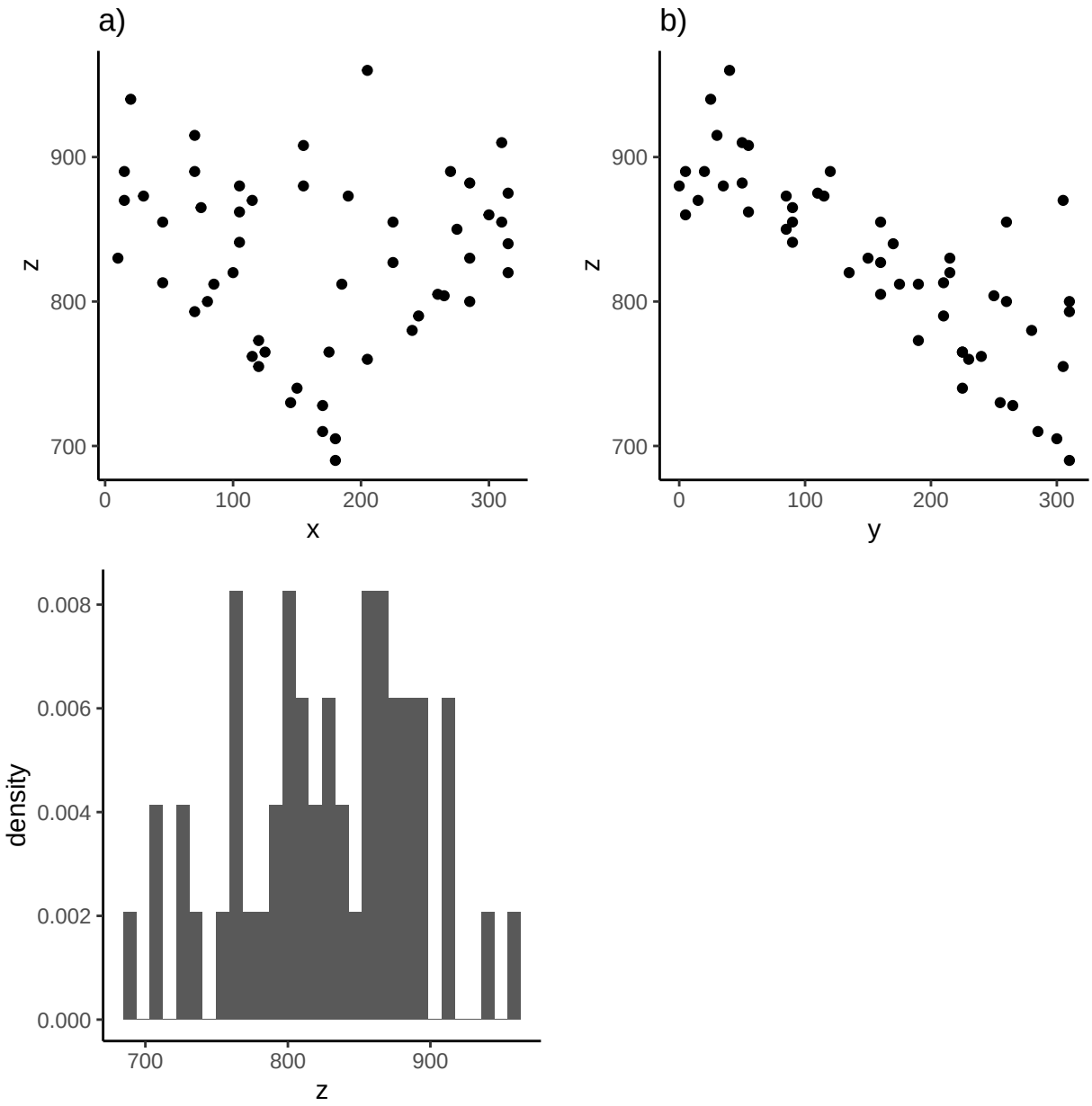


Figure 8: Plot of the data points with respect to their a) x-cordinates and b) y-cordinates; and c) shows the density distribution of the terrain elevation observations.

Problem 2b

Let

$$\{r(\mathbf{x}); \mathbf{x} \in D \subset \mathbb{R}^{\not\vdash}\}$$

be the Gaussian RF that is used to model the domain D.

Given that $E\{r(x)\} = (\mathbf{gx})^T \boldsymbol{\beta}_\mathbf{r}$, $Var\{r(\mathbf{x})\} = \sigma_r^2$ and $Corr(r(\mathbf{x}), r(\mathbf{x'})) = \rho_r(\frac{\tau}{\xi})$. We assume that σ_r^2, ξ are assumed known but $\boldsymbol{\beta}_\mathbf{r}$ is unknown. This is therefore a universal krigging problem.

Let the krigging predictor be:

$$\hat{\mathbf{r}}_\mathbf{0} = \boldsymbol{\alpha}^T \mathbf{r}^\mathbf{d}$$

We discretise the predictor as:

$$\{\mathbf{r}_\boldsymbol{\Delta}(\mathbf{x}) = r(\mathbf{x}) - \mu_r^0 - \sum_{i=1}^{n_g} \beta_r^j g_j(\mathbf{x}); \mathbf{x} \in D\}$$

For the estimator to be unbiased,

$$E\{\hat{\mathbf{r}}_\mathbf{0} - \mathbf{r}_\mathbf{0} = 0\} \implies \sum_{i=1}^{m} \alpha_i E\{r_i^d\} - E\{\mathbf{r}_\mathbf{0}\} = 0$$

$$\sum_{i=1}^{m} \sum_{j=1}^{n_g} \alpha_i \beta_r^j g_j(\mathbf{x}_i) = \sum_j \beta_r^j g_j(\mathbf{x}_0)$$

For the estimator to be unbiased,
$\sum_{i=1}^m \alpha_i g_j(\mathbf{x}_i) = \sum_j \beta_r^j g_j(\mathbf{x}_0)$.

$$\begin{aligned} Var\{\hat{\mathbf{r}}_\mathbf{0} - \mathbf{r}_\mathbf{0}\} &= E\{(\hat{\mathbf{r}}_\mathbf{0} - \mathbf{r}_\mathbf{0})^2\} \\ &= Var\{\alpha_i\{r_i^d\} - \mathbf{r}_\mathbf{0}\} \\ &= \sigma^2 \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i \alpha_j \rho_{ij} + \sigma^2 + 2\sigma^2 \sum_{j=1}^{m} \alpha_j \rho_{j0} \end{aligned}$$

Hence, we find $\hat{\boldsymbol{\alpha}}$ such that

$$\hat{\boldsymbol{\alpha}} = argmin_{\boldsymbol{\alpha}} Var\{\hat{\mathbf{r}}_\mathbf{0} - \mathbf{r}_\mathbf{0}\}$$

and subject to the constraint $\sum_{i=1}^m \alpha_i g_j(\mathbf{x}_i) = \sum_j \beta_r^j g_j(\mathbf{x}_0)$ for $j = 1, 2, ..., n_g$.

Problem 2c

Considering the case with $E(r(\mathbf{x})) = \beta_1$, we estimated the universal krigging predictor and variance as follows:

Problem 2d

The resulting polynomial function becomes:

$$\mathbf{(gx)} = (1, x_v, x_h, x_v x_h, x_v^2, x_h^2)$$

The expected value of $r(\mathbf{x})$ then is:

$$E\{r(\mathbf{x})\} = \beta_1 + \beta_2 x_v + \beta_3 x_h + \beta_4 x_v x_h + \beta_5 x_v^2 + \beta_6 x_h^2.$$

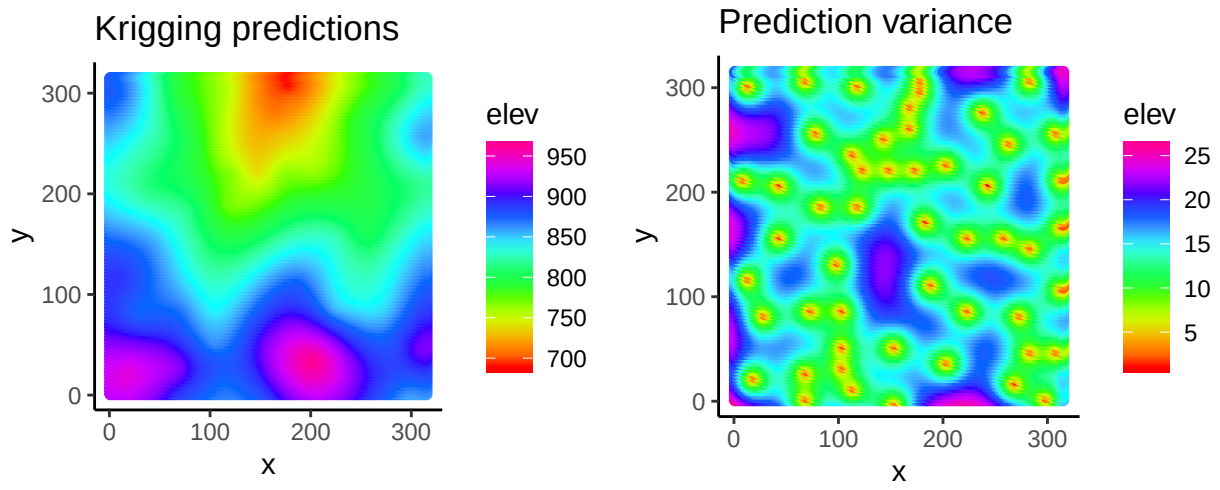
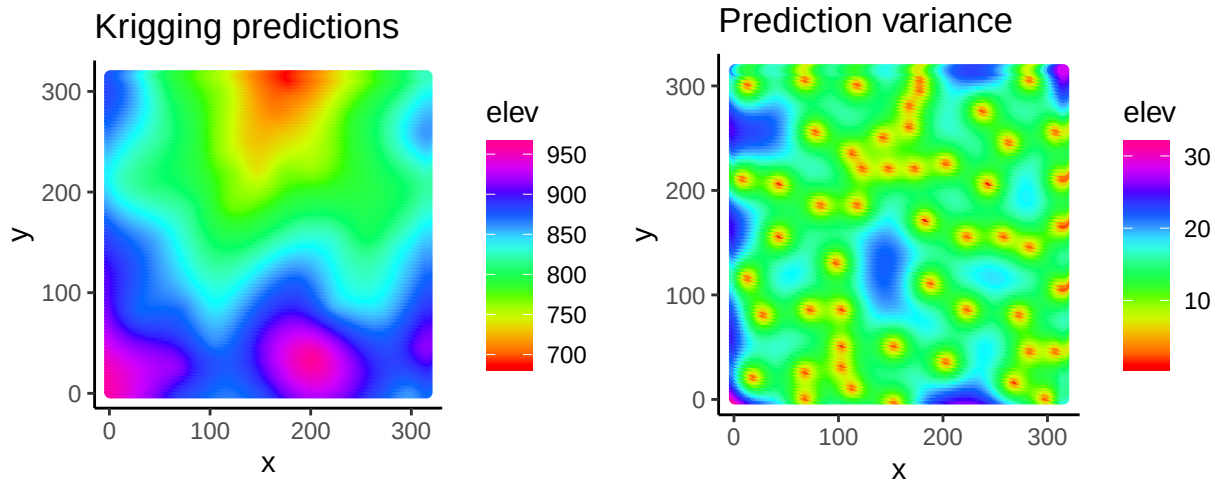


Figure 9: Krigging predictions and prediction variance of the ordinary krigging method.



We present the predictions and the associated variance in Figure **??**

Problem 2e

We now consider the grid node, $\mathbf{x_0} = (100, 100)$. Using the ordinary krigging, we estimated the predicted mean as 838.6781414 and the predicted variance as 10.044233. Assuming Gaussianity of the data,

$$\begin{aligned} P(r\{\mathbf{x_0}\} > 850) &= P\left(\frac{r\{\mathbf{x_0}\} - E(r\{\mathbf{x_0}\})}{\sqrt{Var(r\{\mathbf{x_0}\})}} > \frac{850 - E(r\{\mathbf{x_0}\})}{\sqrt{Var(r\{\mathbf{x_0}\})}}\right) \\ &= 1 - \Phi\left(\frac{850 - E(r\{\mathbf{x_0}\})}{\sqrt{Var(r\{\mathbf{x_0}\})}}\right) \end{aligned}$$

The resulting probability is 0.13.

To obtain the elevation for which it is 0.90 probability that the true elevation is below it, we used the formular,

$$P(r\{\mathbf{x_0}\} > r\{\mathbf{x_{new}}\}) = 0.90$$
$$r\{\mathbf{x_{new}}\} = E(r\{\mathbf{x_0}\}) + \phi(0.90)\sqrt{Var(r\{\mathbf{x_0}\})}$$

We obatained 851.55m to be that elevation that satifies the preamble.

Problem 3

We consider the stationary Gaussian RF $\{r((\mathbf{x}); \mathbf{x} \in D \subset \mathbb{R}^2$. With $D : [(1, 30) \times (1, 30)]$.

With:

$$\begin{aligned} E\{r(\mathbf{x})\} &= \mu_r = 0 \\ Var\{r(\mathbf{x})\} &= \sigma_r^2 \\ Corr\{r(\mathbf{x}), r(\mathbf{x}')\} & \\ &= exp\{-\frac{\tau}{\xi_r}\} \end{aligned} \tag{37}$$

with $\tau = |\mathbf{x} - \mathbf{x}'|$

Problem 3ac

Discretize the random field with $L : [(1, 30) \times (1, 30)] \in D$ with parameters σ_r^2 and $\xi_r = 3$. And simulate realizations.

Create four realizations that are presented in Figure 10

Problem 3bc

The theoretical variogram function can be written as:

$$\gamma_r(\tau) = \sigma_r^2(1 - \rho_r(\tau)) \tag{38}$$

We plot the emprical variogram from our realisations and compare to the theoretical variogram.

The results are plotted in Figure 11. None of the empirical variograms seem to be very accurate. The seem to capture the shape of the theoretical variogram at low distances, however at large distances the empirical variograms vary alot, this is probably due to a low amount of points that have a large distance in between.

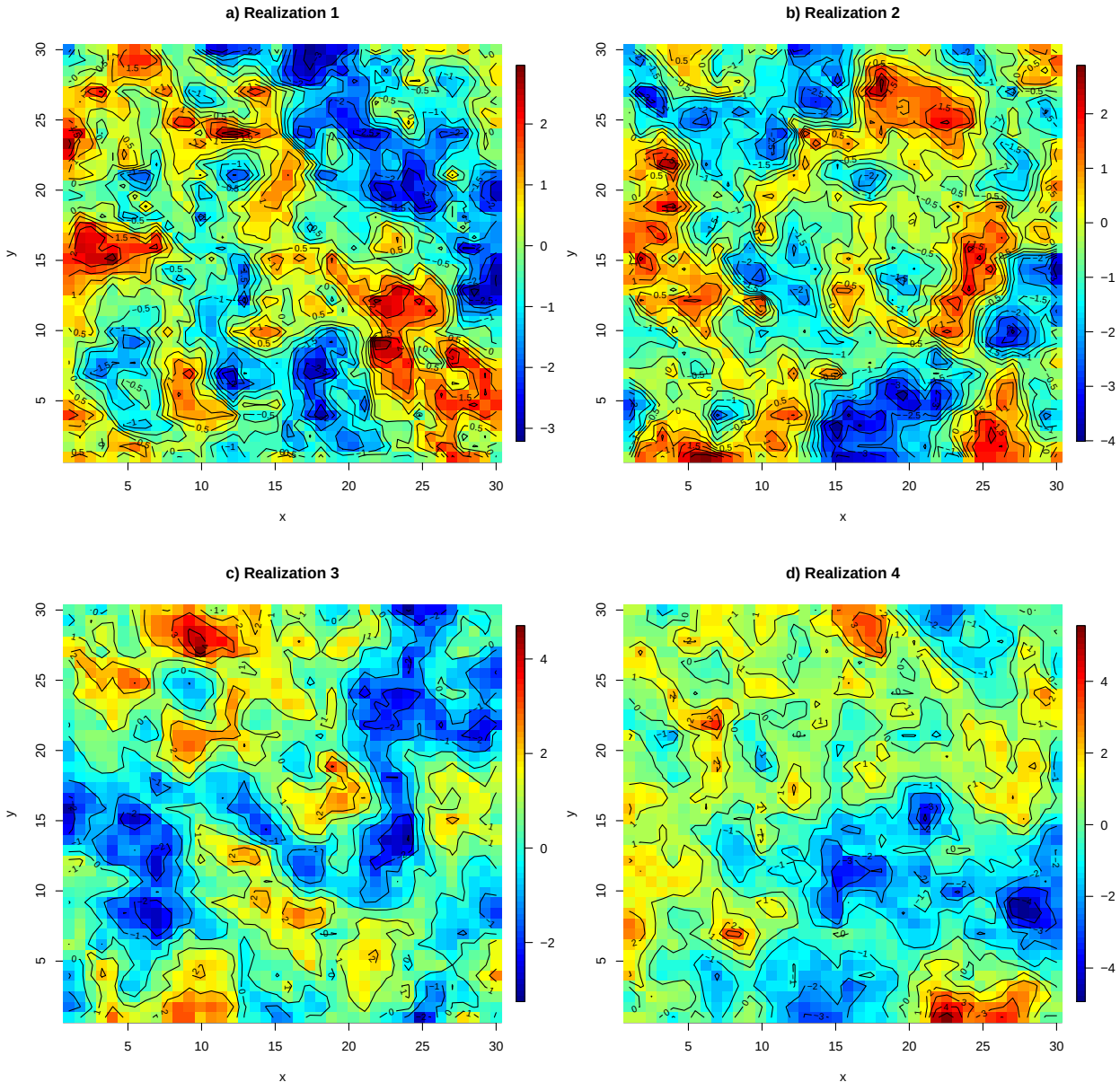


Figure 10: Realizations of field with $\xi = 3$, $\sigma^2 = 2$

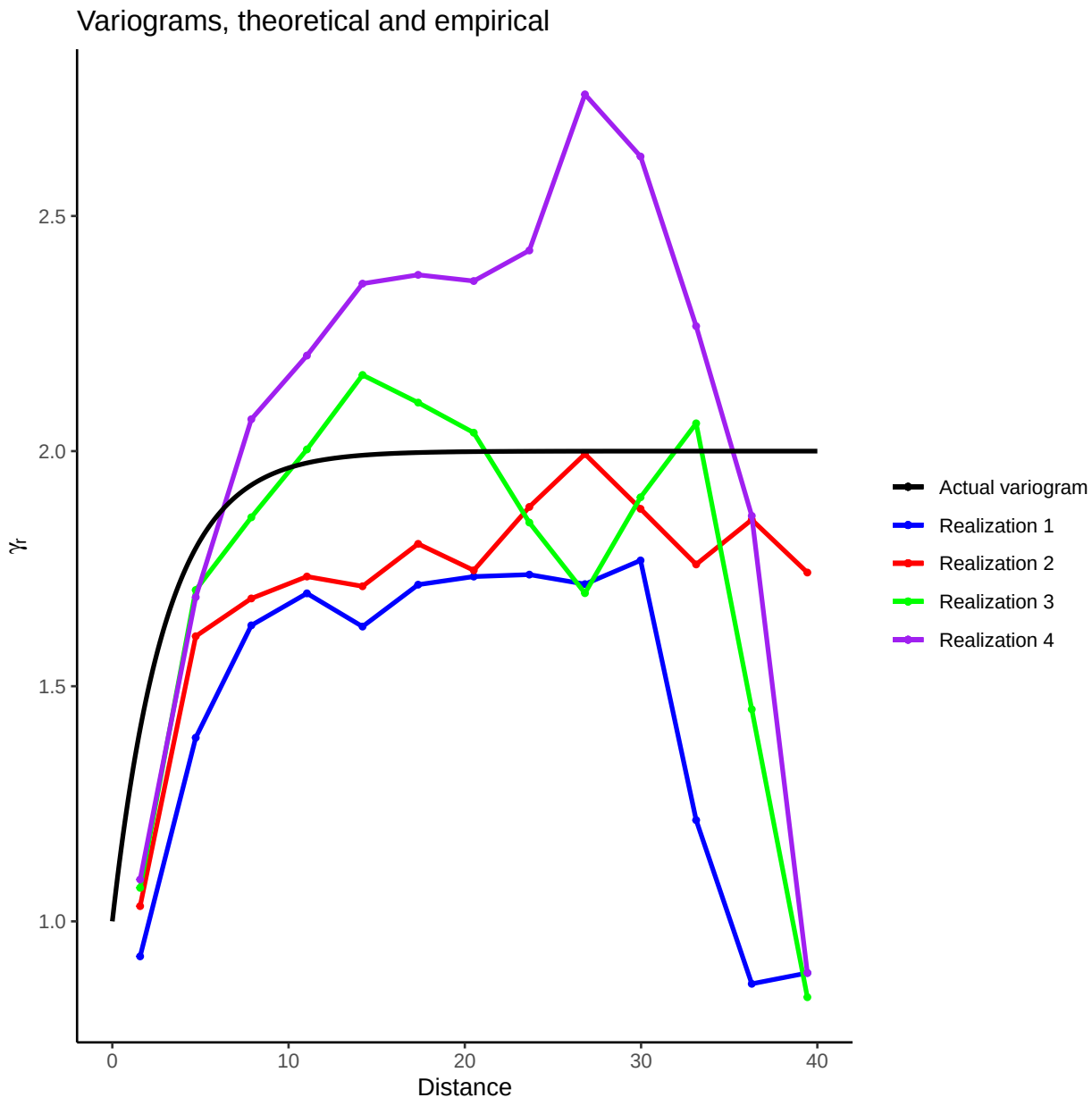


Figure 11: Empirical variogram vs. theoretical variogram of realisation in Figure 10.$\xi = 3$, $\sigma^2 = 2$

Problem 3de

Generate 36 point uniformly on L. And simulate a realisation over those 36 points. Will later take on these as exact observations.

Use **likfit** to estimate the parameters of our field using our 36 exact observations.

The summary of the **likfit** is displayed inbelow:

```
summary(likfit.36)
```

```
## Summary of the parameter estimation
## ----------------------------------
## Estimation method: maximum likelihood
##
## Parameters of the mean component (trend):
##   beta
## 0.8442
##
## Parameters of the spatial component:
##    correlation function: exponential
##       (estimated) variance parameter sigmasq (partial sill) =  2.255
##       (estimated) cor. fct. parameter phi (range parameter)  =  1.869
##    anisotropy parameters:
##       (fixed) anisotropy angle = 0  ( 0 degrees )
##       (fixed) anisotropy ratio = 1
##
## Parameter of the error component:
##       (estimated) nugget =  0
##
## Transformation parameter:
##       (fixed) Box-Cox parameter = 1 (no transformation)
##
## Practical Range with cor=0.05 for asymptotic range: 5.599689
##
## Maximised Likelihood:
##    log.L n.params      AIC      BIC
## "-64.19"      "4"  "136.4"  "142.7"
##
## non spatial model:
##    log.L n.params      AIC      BIC
## "-67.12"      "2"  "138.2"  "141.4"
##
## Call:
## likfit(geodata = geo, ini.cov.pars = c(1, 1), cov.model = "exponential")
```

Using the 36 point we estimate $\hat{\sigma}^2 = 2.254746$ and $\hat{\xi} = 1.8692279$. We repeat the process in 3d for $n = 6$, $n = 64$, $n = 100$.

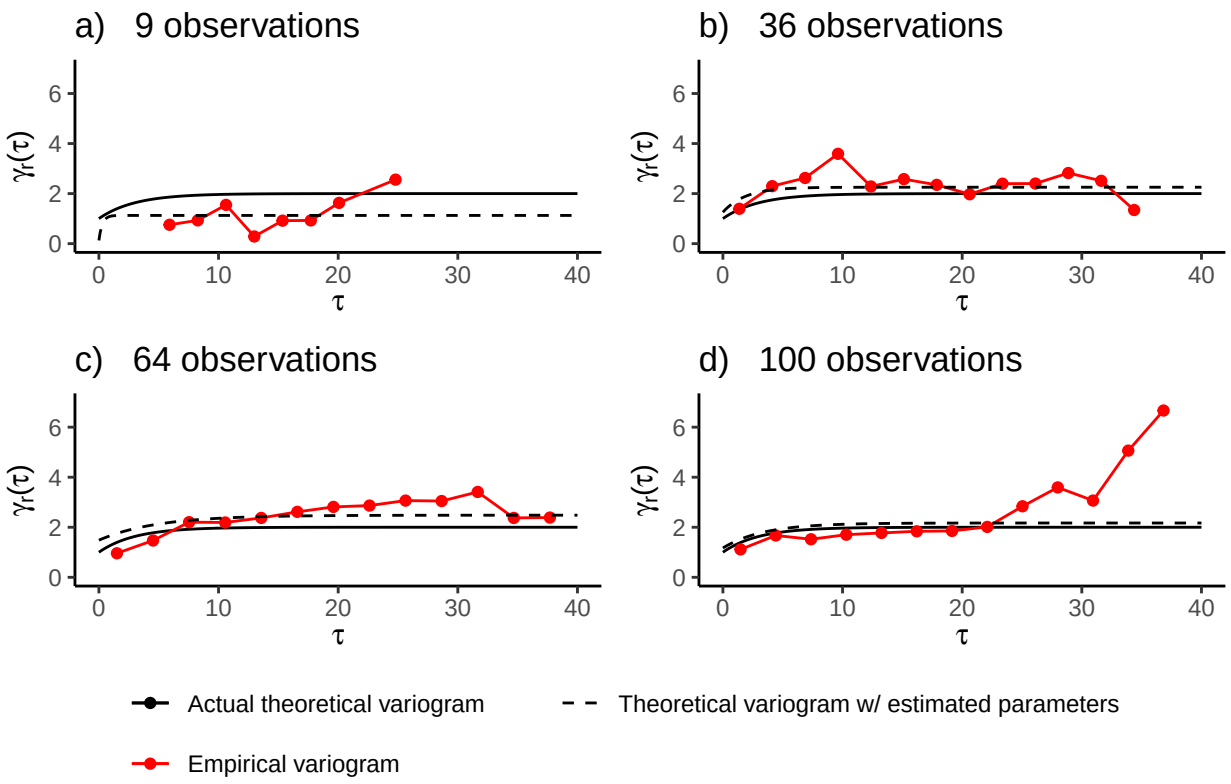


Figure 12: Variograms from uniform drawn points $\xi = 3$, $\sigma^2 = 2$, maximum likelihood fit

The ML estimates of the parameters were:

n	$\hat{\sigma}^2_{ML}$	$\hat{\xi}_{ML}$
9	1.1269907	0.28203
36	2.254746	1.8692279
64	2.4823404	4.7979249
100	2.1700308	3.3909212
Actual	2	3

We see that the estimates becomes quite accurate at large values of n, and be very off at lower values of n

We plot the empirical variograms vs theoretical.

These variograms are displayed in Figure 12. They seem to vary qutite much, but that can probably be blamed on a low amount of observation points. As we draw a very limited amount of points, some distances will be less represented than others, which will cause large variations in estimation. This can especially be seen for $n = 9$

Problem 3f

In this problem we have used maxmium likelihood to estiamte the parameters of the correlation funciton. We have found that using maximum likelihood to estiamte parameters seems to be relatively fast using the number of parameters and observations we have here (although **likfit** warns that the opposite might be the case). With enough data points maximum likelihood seems to be quite good at finding the correct parameters.