554-2022-Assignment2_key

Question 1

a. Maps of the observed counts $Y_i's$

b. Maps of the expected counts $E_i's$, and

c. Maps of the SMRS defined as SMR $= \frac{Y_i}{E_i}$.

```r
### load libraries and data
library(RColorBrewer)
library(rgdal)
library(INLA)
load("hw data/HW2data.Rdata")
### lung cancer data
# load("ObsTrivariate-mod.Rdata")
# load("ExpTrivariate.Rdata")

### calculate the SMR
SMR.lung = Obs.mv3[,"Lung"]/Exp.mv3[,"Lung"]
ColorSMR <- brewer.pal(7, "BrBG")[7:1]
ColorProb <- brewer.pal(7, "RdYlGn")[7:1]
par(mfrow = c(1, 3))
par(mar = c(1, 1, 2, 1) + 0.1)
plot(VR.cart, col = ColorProb[as.numeric(cut(Obs.mv3[,"Lung"],
    breaks = quantile(Obs.mv3[,"Lung"],probs =
                        seq(0,1,1/7)),include.lowest=TRUE))])
title("Observed", cex = 0.75)
legend(x = "bottomright", fill = ColorProb[7:1],
       legend = c(">107", "40-107", "19-40", "10-19",
                   "6-10", "3-6", "<3"), cex = 0.65,
       inset = 0.03, title = "Observed")

plot(VR.cart, col = ColorProb[as.numeric(cut(Exp.mv3[,"Lung"],
    breaks = quantile(Obs.mv3[,"Lung"],probs =
                        seq(0,1,1/7)), include.lowest=TRUE))])
title("Expected", cex = 0.75)
legend(x = "bottomright", fill = ColorProb[7:1],
       legend = c(">107", "40-107", "19-40", "10-19",
                   "6-10", "3-6", "<3"), cex = 0.65,
       inset = 0.03, title = "Expected")

plot(VR.cart, col = ColorSMR[as.numeric(
  cut(SMR.lung, c(-0.1, 1/1.5, 1/1.25, 1/1.1,1.1, 1.25, 1.5, 100)))])
title("SMR", cex = 0.75)
```
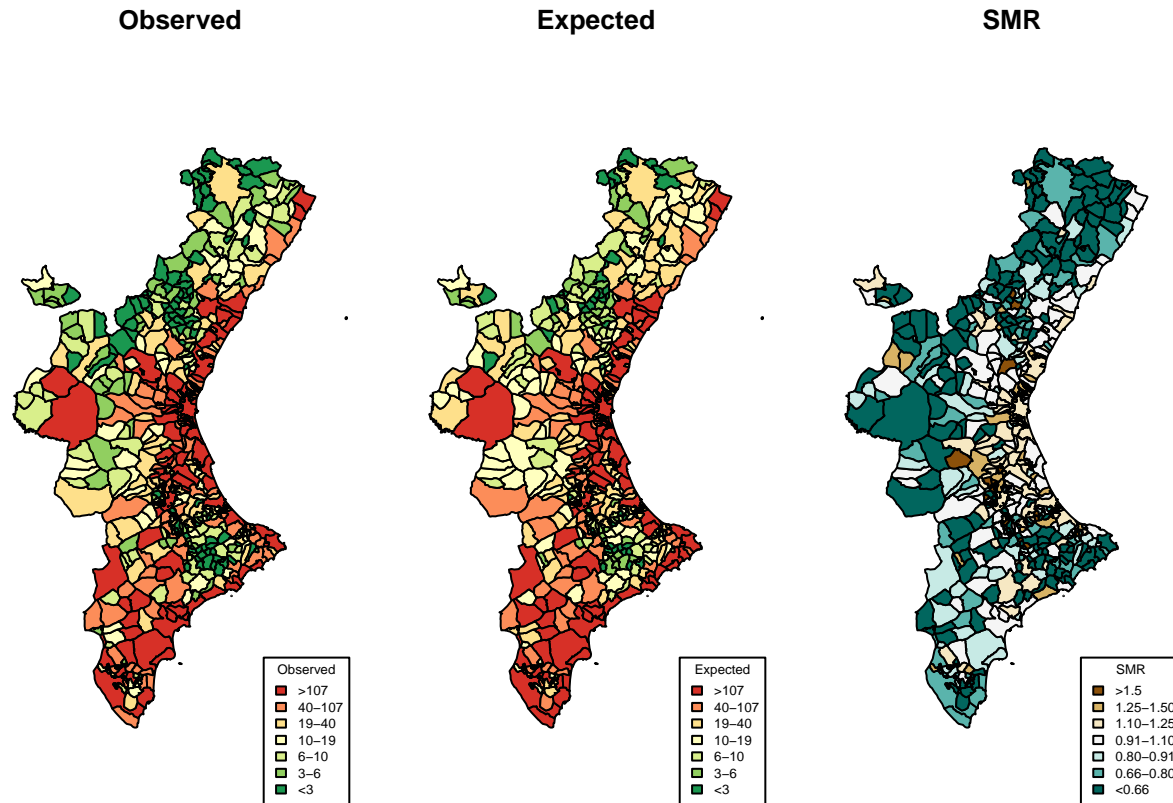
```
legend(x = "bottomright", fill = ColorSMR[7:1],
       legend = c(">1.5", "1.25-1.50", "1.10-1.25",
                  "0.91-1.10", "0.80-0.91", "0.66-0.80","<0.66"),
       cex = 0.65, inset = 0.03, title = "SMR")
```

Observed **Expected** **SMR**

The unsmoothed SMRs are quite variable–areas in the eastern gulf of the Valencian Region exhibit higher risks of lung cancer whereas the inland areas have relatively low risks. We can also observe some spatial dependence between neighboring areas as they exhibit similar SMRs.

d. SMRs verus the estimated standard errors.

```
par(pty="s")
plot(sqrt(SMR.lung/Exp.mv3[,"Lung"])~SMR.lung,
     ylim = c(0,1.5), xlim=c(0,1.5),xlab="SMR",ylab="Estimated standard error",
     main="")
abline(coef = c(0,1),col="grey")
```
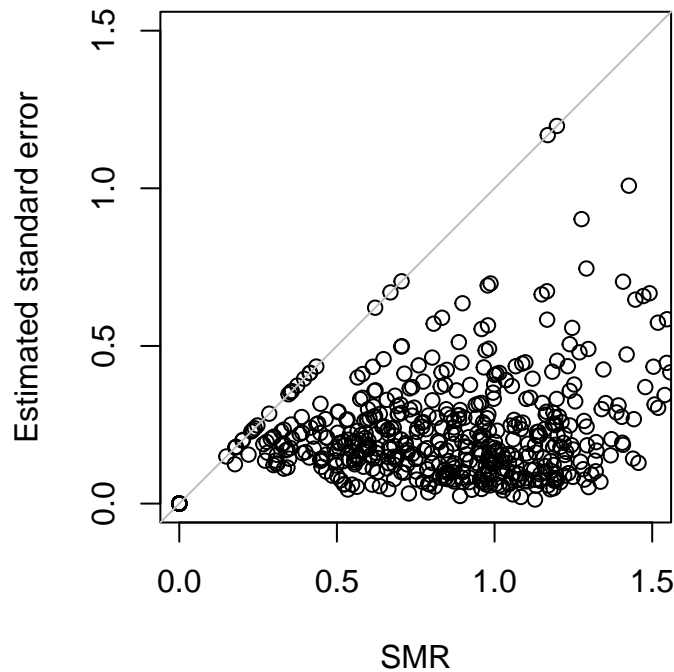
Question 2: Smooth SMRs using Poisson-Lognormal model

```
lung.data = data.frame(O = Obs.mv3[,"Lung"], E = Exp.mv3[,"Lung"], id.node = 1:nrow(Obs.mv3))
result.INLA <- inla(O ~ 1 + f(id.node, model="iid"),data=lung.data,
                     family="poisson", E=E,
                     control.predictor = list(compute = TRUE),
                     control.compute= list(return.marginals=TRUE))
```

a. The posterior medians and 95% intervals for β_0 and σ_e

```
### beta: result.INLA$summary.fixed
### sigm^2a_e: result.INLA$summary.hyper
result  <-  data.frame("posterior.Median"=c(result.INLA$summary.fixed$`0.5quant`,
                                            1/sqrt(result.INLA$summary.hyper$`0.5quant`)),
                "lower.CI"=c(result.INLA$summary.fixed$`0.025quant`,
                             1/sqrt(result.INLA$summary.hyper$`0.975quant`)),
                "upper.CI"=c(result.INLA$summary.fixed$`0.975quant`,
                             1/sqrt(result.INLA$summary.hyper$`0.025quant`)))
rownames(result) = c("beta","sigma")
knitr::kable(round(result,2), "simple")
```

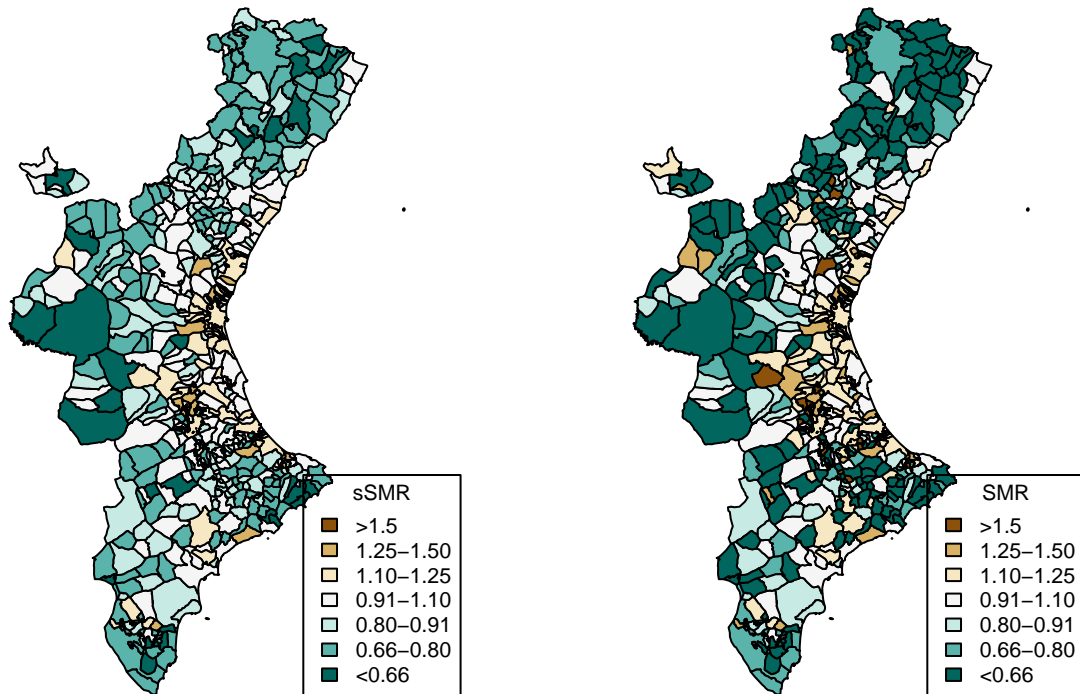|	posterior.Median	lower.CI	upper.CI
beta	-0.13	-0.17	-0.10
sigma	0.25	0.23	0.28

b. Map of relative risk estimates vs SMR.

```
### relative risks
par(mfrow = c(1, 2))
par(mar = c(1, 1, 2, 1) + 0.1)
plot(VR.cart, col = ColorSMR[as.numeric(
  cut(result.INLA$summary.fitted.values$`0.5quant`,
      c(-0.1, 1/1.5, 1/1.25, 1/1.1,1.1, 1.25, 1.5, 100)))])
title("Poisson-Lognormal non-spatial", cex = 0.75)
legend(x = "bottomright", fill = ColorSMR[7:1],
       legend = c(">1.5", "1.25-1.50", "1.10-1.25",
                  "0.91-1.10", "0.80-0.91", "0.66-0.80","<0.66"),
       cex = 0.65, inset = 0.03, title = "sSMR")

plot(VR.cart, col = ColorSMR[as.numeric(
  cut(SMR.lung,c(-0.1, 1/1.5, 1/1.25, 1/1.1,1.1, 1.25, 1.5, 100)))])
title("SMR", cex = 0.75)
legend(x = "bottomright", fill = ColorSMR[7:1],
       legend = c(">1.5", "1.25-1.50", "1.10-1.25",
                  "0.91-1.10", "0.80-0.91", "0.66-0.80","<0.66"),
       cex = 0.65, inset = 0.03, title = "SMR")
```


c. Comparison of relative risk estimates to SMR

The posterior RR estimates are smoothed towards the center as expected. Areas with high SMR's are shrinked to lower values and areas with zero observed counts are pulled toward the center. RR estimates also exhibit lower variablility compared to SMR's.

```
par(pty="s")
plot(result.INLA$summary.fitted.values$`0.5quant`~SMR.lung,
     ylim = c(0,1.5), xlim=c(0,1.5),xlab="SMR",ylab="Poisson-Lognormal",
     main="Comparision of relative risk estimates")
abline(coef = c(0,1),col="grey")
```

Comparision of relative risk estimates
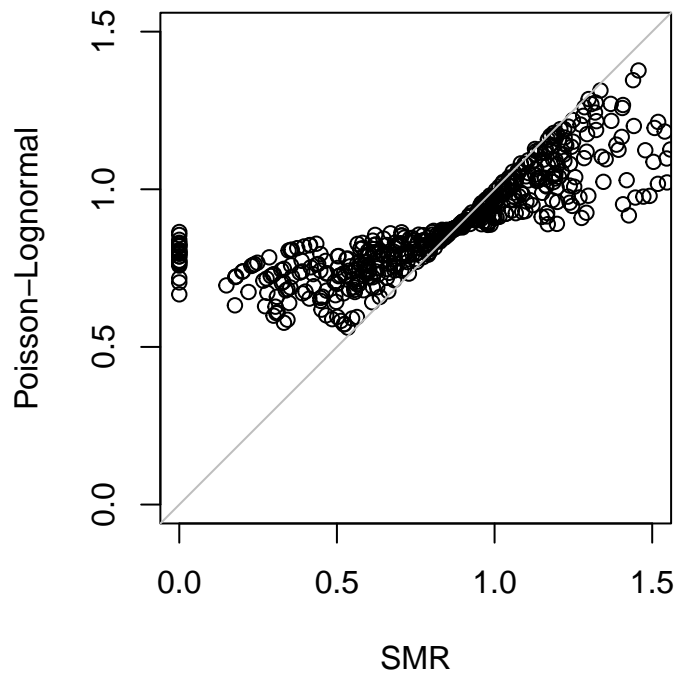
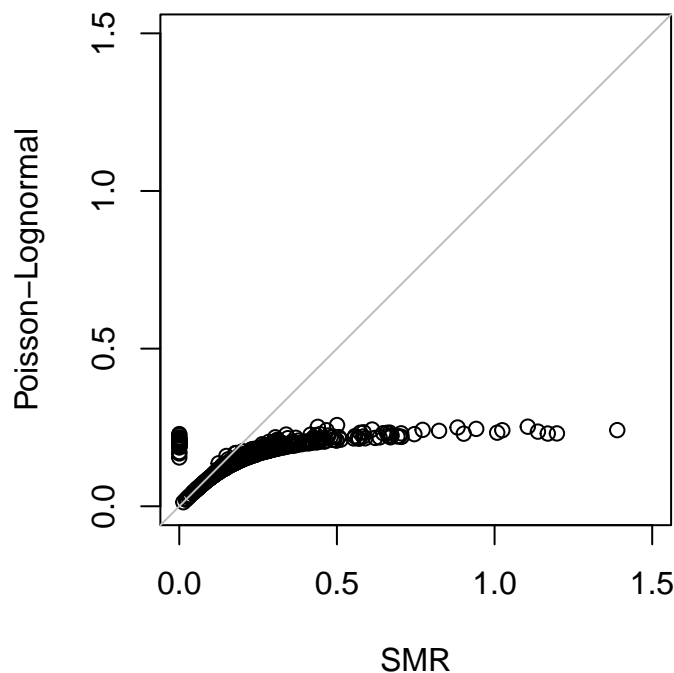**d. Posterior standard deviations of the RRs against the standard deviations of SMR's.**

The posterior standard deviations are in general smaller than the standard deviations of SMR's due to smoothing. Areas with zero observed counts are exceptions since the standard deviations of SMR's are zero.

```
par(pty="s")
plot(result.INLA$summary.fitted.values$sd~sqrt(SMR.lung/Exp.mv3[,"Lung"]),
     ylim = c(0,1.5), xlim=c(0,1.5),xlab="SMR",ylab="Poisson-Lognormal",
     main="Comparision of relative risk estimates")
abline(coef = c(0,1),col="grey")
```

Comparision of relative risk estimates

Question 3: Smooth SMRs using Poisson-Lognormal-Spatial model

```r
# formula <- O ~ 1 + f(id.node, model="bym2",graph = "~/Downloads/VR.graph",
#                                scale.model=T, constr=T,
#                                hyper=list(phi=list(prior="pc", param=c(0.5,0.5),initial=1),prec=list(j

formula <- O ~ 1 + f(id.node, model="bym2",graph = "hw data/VR.graph",
                             scale.model=T, constr=T,
                             hyper=list(phi=list(prior="pc", param=c(0.5,0.5),initial=1),prec=list(pr

result.INLA2 <- inla(formula,data=lung.data,
                     family="poisson", E=E,
                     control.predictor = list(compute = TRUE),
                     control.compute=list(return.marginals=TRUE))
```

a. The total variance of the random effects is **0.07**, and a total variance **0.96** coming from the spatial random effect.

```r
result2  <- data.frame("posterior.Median"=c(result.INLA2$summary.fixed$`0.5quant`,
                                            1/(result.INLA2$summary.hyper$`0.5quant`[1]),
                                            result.INLA2$summary.hyper$`0.5quant`[2]),
                 "lower.CI"=c(result.INLA2$summary.fixed$`0.025quant`,
                             1/(result.INLA2$summary.hyper$`0.975quant`[1]),
                             result.INLA2$summary.hyper$`0.025quant`[2]),
                 "upper.CI"=c(result.INLA2$summary.fixed$`0.975quant`,
```

```
                               1/(result.INLA2$summary.hyper$`0.025quant`[1]),
                               result.INLA2$summary.hyper$`0.975quant`[2]))
rownames(result2) = c("beta","total_var","prop_spatial")
knitr::kable(round(result2,2), "simple")
```

	posterior.Median	lower.CI	upper.CI
beta	-0.15	-0.18	-0.13
total_var	0.07	0.05	0.09
prop_spatial	0.96	0.86	1.00

b. Map of relative risk estimates and comparison between SMR's and non-spatial RR's.

```
### relative risks
par(mfrow = c(1, 2))
par(mar = c(1, 1, 2, 1) + 0.1)
plot(VR.cart, col = ColorSMR[as.numeric(
  cut(result.INLA2$summary.fitted.values$`0.5quant`,
      c(-0.1, 1/1.5, 1/1.25, 1/1.1,1.1, 1.25, 1.5, 100)))])
title("Poisson-Lognormal spatial", cex = 0.75)
legend(x = "bottomright", fill = ColorSMR[7:1],
       legend = c(">1.5", "1.25-1.50", "1.10-1.25",
                  "0.91-1.10", "0.80-0.91", "0.66-0.80","<0.66"),
       cex = 0.65, inset = 0.03, title = "sSMR")

plot(VR.cart, col = ColorSMR[as.numeric(
  cut(result.INLA$summary.fitted.values$`0.5quant`,
      c(-0.1, 1/1.5, 1/1.25, 1/1.1,1.1, 1.25, 1.5, 100)))])
title("Poisson-Lognormal non-spatial", cex = 0.75)
legend(x = "bottomright", fill = ColorSMR[7:1],
       legend = c(">1.5", "1.25-1.50", "1.10-1.25",
                  "0.91-1.10", "0.80-0.91", "0.66-0.80","<0.66"),
       cex = 0.65, inset = 0.03, title = "sSMR")
```
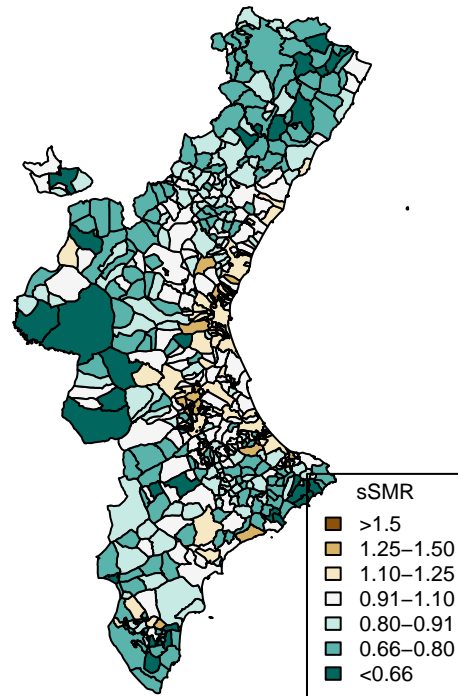
# Poisson–Lognormal spatial
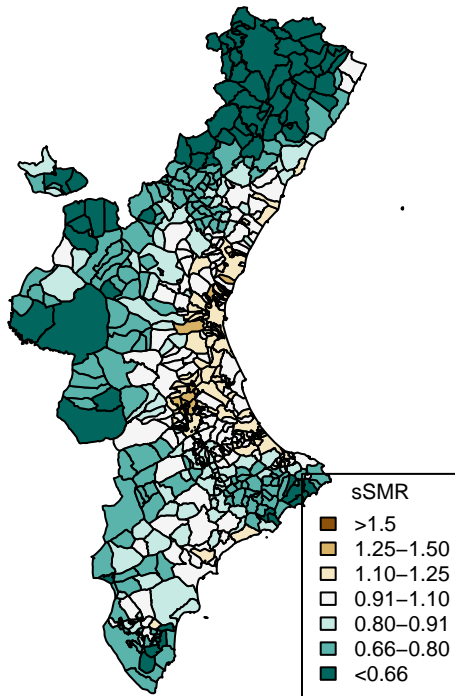
Poisson–Lognormal non–spatial

The BYM model picks up highly dependent geographical patterns (i.e.northwestern region). Dependence between spatial units is evident, with the SMRs of nearby municipalities being very similar in general.

The non-spatial RR estimates exhibit less spatial dependence.

```
par(mfrow = c(1, 2))
par(pty="s")
plot(result.INLA2$summary.fitted.values$`0.5quant`~result.INLA$summary.fitted.values$`0.5quant`,
     ylim = c(0,1.5), xlim=c(0,1.5),
     xlab="nonspatial Poisson-Lognormal",
     ylab="spatial Poisson-Lognormal",
     main="")
abline(coef = c(0,1),col="grey")

par(pty="s")
plot(result.INLA2$summary.fitted.values$`0.5quant`~SMR.lung,
     ylim = c(0,1.5), xlim=c(0,1.5),
     xlab="SMR",
     ylab="spatial Poisson-Lognormal",
     main="")
abline(coef = c(0,1),col="grey")
```
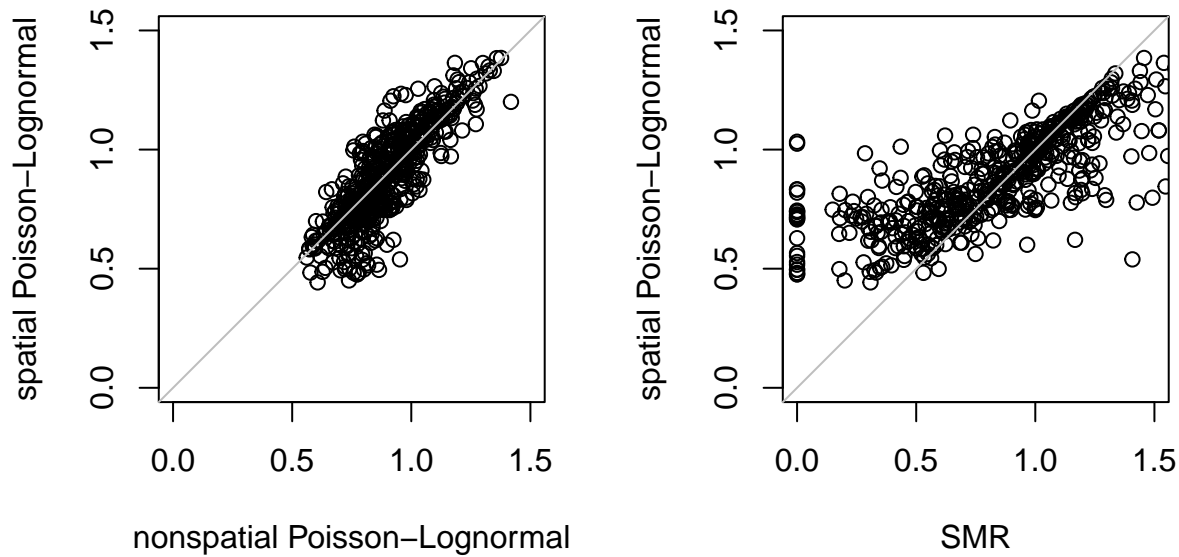
The smoothed RR estimates from the two models are similar to each other, and both are shrinked towards the center compared to the raw SMR's.

Question 4

```r
# small correction for if any observations are 0: add 0.5 to both Y and E
Obs.fixed <- ifelse(Obs.mv3[,"Lung"] == 0, 0.5, Obs.mv3[,"Lung"])
Exp.fixed <- ifelse(Exp.mv3[,"Lung"] == 0, 0.5, Exp.mv3[,"Lung"])
SMR.lung = Obs.fixed / Exp.fixed
sigma_e <- sqrt(SMR.lung/Exp.fixed)

lung.data = data.frame(O = Obs.fixed, E = Exp.fixed, id.node = 1:nrow(Obs.mv3))

lung.data$log_theta_hat <-  log(SMR.lung) # Z_i
lung.data$v <- sigma_e^2/SMR.lung^2

formula3  <- log_theta_hat ~ 1 + f(id.node, model = "iid")


result.INLA3 <- inla(formula3,
                      data=lung.data,
                      family="gaussian",
                      control.family = list(hyper = list(prec = list(initial= log(1), fixed = TRUE))),
                      scale = 1/lung.data$v,
                      control.predictor = list(compute = TRUE),
                      control.compute= list(return.marginals=TRUE))


### relative risks
par(mfrow = c(1, 2))
par(mar = c(1, 1, 2, 1) + 0.1)
plot(VR.cart, col = ColorSMR[as.numeric(
  cut(exp(result.INLA3$summary.fitted.values$`0.5quant`),
      c(-0.1, 1/1.5, 1/1.25, 1/1.1,1.1, 1.25, 1.5, 100)))])
```

```
title("Lognormal", cex = 0.75)
legend(x = "bottomright", fill = ColorSMR[7:1],
        legend = c(">1.5", "1.25-1.50", "1.10-1.25",
                   "0.91-1.10", "0.80-0.91", "0.66-0.80","<0.66"),
        cex = 0.65, inset = 0.03, title = "sSMR")

plot(VR.cart, col = ColorSMR[as.numeric(
  cut(result.INLA$summary.fitted.values$`0.5quant`,
      c(-0.1, 1/1.5, 1/1.25, 1/1.1,1.1, 1.25, 1.5, 100)))])
title("Poisson-Lognormal non-spatial", cex = 0.75)
legend(x = "bottomright", fill = ColorSMR[7:1],
        legend = c(">1.5", "1.25-1.50", "1.10-1.25",
                   "0.91-1.10", "0.80-0.91", "0.66-0.80","<0.66"),
        cex = 0.65, inset = 0.03, title = "sSMR")
```
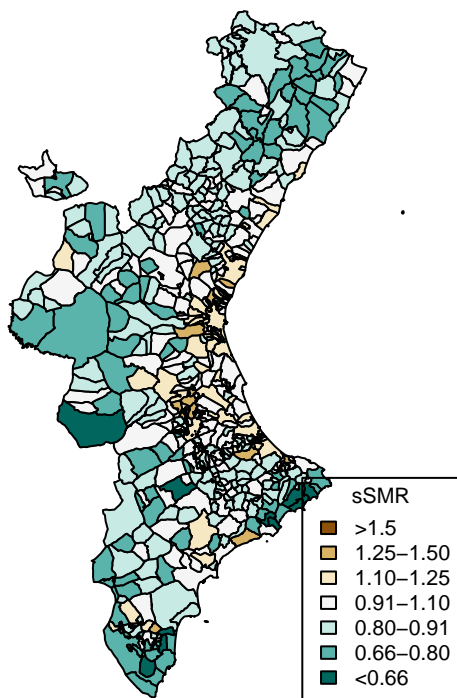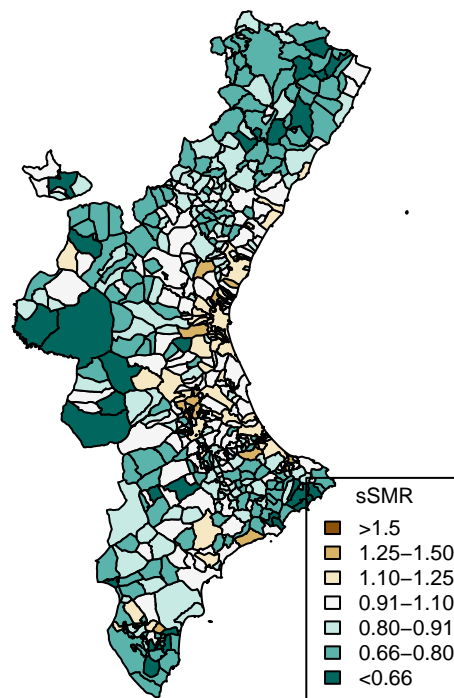
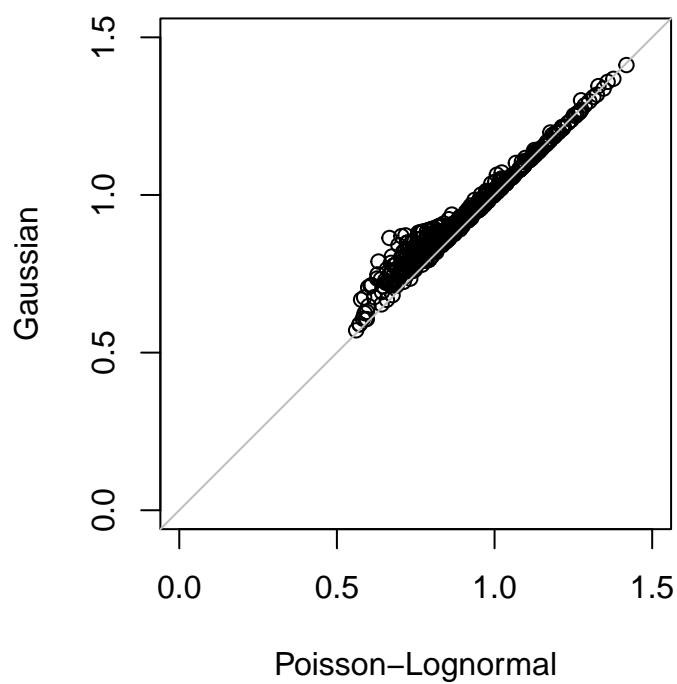## Lognormal   Poisson–Lognormal non–spatial


```
par(pty="s")
plot(exp(result.INLA3$summary.fitted.values$`0.5quant`)~result.INLA$summary.fitted.values$`0.5quant`,
     ylim = c(0,1.5), xlim=c(0,1.5),
     xlab="Poisson-Lognormal",
     ylab="Gaussian",
     main = "Comparision of relative risk estimates ")
abline(coef = c(0,1),col="grey")
```

Comparision of relative risk estimates

The relative risks from the Poisson-Lognormal model and the model fit in this question are incredibly similar, as evidenced by the points above falling very close to the line with slope = 1, intercept = 0.