5. sklop: Normalni model z dvema parametroma

Nina Ruzic Gorenjec

1 Primer

Podan imamo naslednji vzorec visin (metri) studentov moskega spola:

```
x <- c(1.91, 1.94, 1.68, 1.75, 1.81, 1.83, 1.91, 1.95, 1.77, 1.98,
       1.81, 1.75, 1.89, 1.89, 1.83, 1.89, 1.99, 1.65, 1.82, 1.65,
       1.73, 1.73, 1.88, 1.81, 1.84, 1.83, 1.84, 1.72, 1.91, 1.63)
```

Zanima nas povprecna visina studentov, kjer standardnega odklona ne poznamo.

2 Verjetnostni model za nas primer

Vzorec $X_1, X_2, \ldots, X_n$, kjer je:

- $n = 30$ stevilo studentov,
- X_i predstavlja visino i-tega studenta,
- $X_i \mid \mu, \sigma^2 \sim N(\mu, \sigma^2)$,
- $f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

Prostor parametrov je dvorazsezen: $\theta = (\mu, \sigma^2) \in \mathbb{R} \times (0, \infty)$.

3 Ocenjevanje v Bayesovi statistiki

Bayesova formula:

$$\pi(\mu, \sigma^2 \mid x) \propto L(\mu, \sigma^2 \mid x)\, \pi(\mu, \sigma^2).$$

Nastaviti moramo torej dvorazsezno apriorno porazdelitev in dobiti dvorazsezno aposteriorno porazdelitev.

3.1 Verjetje

$$L(\mu, \sigma^2 \mid x) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

3.2 Apriorna porazdelitev

Vedno velja (pogojna verjetnost):

$$\pi(\mu, \sigma^2) = \pi(\mu \mid \sigma^2) \cdot \pi(\sigma^2).$$

3.2.1 Konjugirana apriorna porazdelitev

Za $\mu \mid \sigma^2$ vzamemo

$$\mu \mid \sigma^2 \sim N(\mu_0, \sigma^2/\kappa_0).$$

Porazdelitev σ^2 lahko zapisemo na tri ekvivalentne nacine:

$$\sigma^2 \sim \text{Inv-Gama}(\nu_0/2,\ \nu_0\,\sigma_0^2/2),$$

kar pomeni ravno

$$1/\sigma^2 \sim \text{Gama}(\nu_0/2,\ \nu_0\,\sigma_0^2/2),$$

kar je ekvivalentno

$$\nu_0\,\sigma_0^2/\sigma^2 \sim \text{Gama}(\nu_0/2,\ 1/2) = \chi^2_{\nu_0}.$$

Ideja za parametre apriorne porazdelitve:

- μ_0 je nase prepricanje o μ
- σ_0^2 je nase prepricanje o σ^2 (ker je ravno pricakovana vrednost)
- vecji kot je ν_0 oz. κ_0, bolj verjamete, da je μ enak μ_0 oz. σ^2 enak σ_0^2
- κ_0 oz. ν_0 predstavljata velikost vzorca, na podlagi katerega je vase prepricanje o μ oz. σ^2 osnovano

3.2.2 Neinformativna apriorna porazdelitev

Privzamemo neodvisnost apriornih porazdelitev parametrov:

$$\pi(\mu, \sigma^2) = \pi(\mu) \cdot \pi(\sigma^2).$$

Jeffreyeva apriorna porazdelitev za μ v normalnem modelu z znanim σ^2:

$$\pi(\mu) \propto \sqrt{1/\sigma^2} \propto 1.$$

Jeffreyeva apriorna porazdelitev za σ^2 v normalnem modelu z znanim μ:

$$\pi(\sigma^2) \propto 1/\sigma^2.$$

Dobimo torej

$$\pi(\mu, \sigma^2) \propto 1/\sigma^2,$$

kar je seveda *improper* porazdelitev.

3.3 Aposteriorna porazdelitev

Tudi tu vedno velja (pogojna verjetnost):

$$\pi(\mu, \sigma^2 \mid x) = \pi(\mu \mid \sigma^2, x) \cdot \pi(\sigma^2 \mid x).$$

3.3.1 Aposteriorna porazdelitev pri konjugirani apriorni porazdelitvi

Dobimo:

$$\mu \mid \sigma^2, x \sim N(\mu_n, \sigma^2/\kappa_n),$$

kjer je

$$\kappa_n = \kappa_0 + n,$$
$$\mu_n = \frac{\kappa_0}{\kappa_0 + n}\mu_0 + \frac{n}{\kappa_0 + n}\bar{x}.$$

Aposteriorno porazdelitev σ^2 lahko zapisemo na tri ekvivalentne nacine:

$$\sigma^2 \mid x \sim \text{Inv-Gama}(\nu_n/2,\ \nu_n\,\sigma_n^2/2) \iff 1/\sigma^2 \mid x \sim \text{Gama}(\nu_n/2,\ \nu_n\,\sigma_n^2/2) \iff$$
$$\iff \nu_n\,\sigma_n^2/\sigma^2 \mid x \sim \text{Gama}(\nu_n/2,\ 1/2) = \chi^2_{\nu_n},$$

kjer je

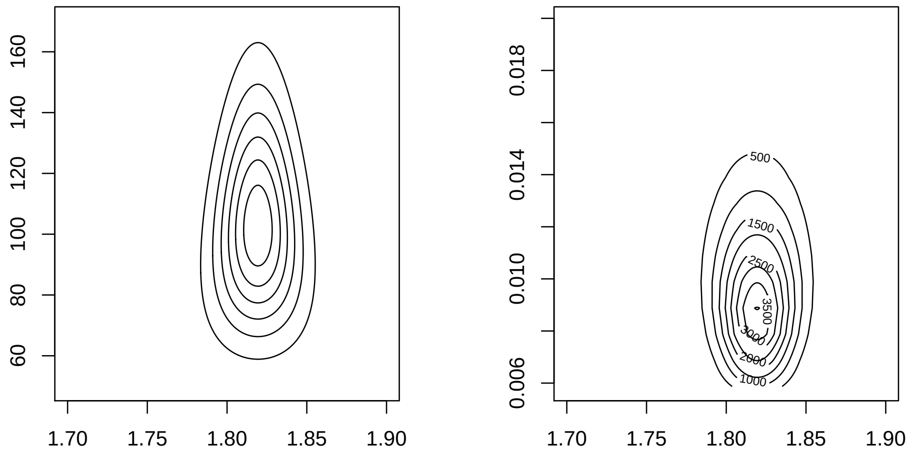
$$\nu_n = \nu_0 + n,$$
$$\nu_n\,\sigma_n^2 = \nu_0\,\sigma_0^2 + (n-1)s^2 + \frac{n\kappa_0}{\kappa_0 + n}(\bar{x} - \mu_0)^2.$$

Ideja:

- Aposteriorna pricakovana vrednost μ_n je utezeno povprecje apriorne pricakovane vrednosti μ_0 in vzorcnega povprecja $\bar{x}$, kjer preko κ_0 kontroliramo, kako mocno verjamemo apriorni pricakovani vrednosti, medtem ko n doloca, kako mocno verjamemo vzorcu.
- Na ravni vsote kvadratov odstopanj je aposteriorna enaka vsoti apriorne, vzorcne in zadnjega clena, ki je posledica pogojne odvisnosti μ od σ^2.

```r
### Izberemo parametre apriorne porazdelitve
mu.0 <- 1.78
kappa.0 <- 1

sigma2.0 <- 0.1^2
nu.0 <- 1

### Izracunamo parametre aposteriorne porazdelitve
n <- length(x)

kappa.n <- kappa.0 + n
mu.n <- kappa.0/kappa.n * mu.0 + n/kappa.n * mean(x)

nu.n <- nu.0 + n
```

```r
sigma2.n <- ( nu.0*sigma2.0 + (n-1)*var(x) +
                n*kappa.0/kappa.n * (mean(x)-mu.0)^2 ) / nu.n

### Narisemo porazdelitev
# Prvi nacin: s parametrom 1/sigma^2
mu <- seq(1.7, 1.9, 0.0001)
prec2 <- seq(50, 170, 1)

apost <- matrix(NA, nrow = length(mu), ncol = length(prec2))
for (i in 1:length(mu)) {
  for (j in 1:length(prec2)) {
    apost[i, j] = dnorm(mu[i], mean = mu.n, sd = sqrt(1/(kappa.n*prec2[j]))) *
      dgamma(prec2[j], nu.n/2, nu.n*sigma2.n/2)
  }
}

par(mfrow = c(1,2))
contour(mu, prec2, apost)

# Drugi nacin: s parametrom sigma^2
#install.packages("invgamma")
library(invgamma)
mu <- seq(1.7, 1.9, 0.0001)
sigma2 <- seq(1/170, 1/50, 0.001)

apost <- matrix(NA, nrow = length(mu), ncol = length(sigma2))
for (i in 1:length(mu)) {
  for (j in 1:length(sigma2)) {
    apost[i, j] = dnorm(mu[i], mean = mu.n, sd = sqrt(sigma2[j]/kappa.n)) *
      dinvgamma(sigma2[j], shape=nu.n/2, rate=nu.n*sigma2.n/2)
  }
}

contour(mu,sigma2,apost)
```




```r
par(mfrow=c(1,1))
```

3.3.2 Aposteriorna porazdelitev pri neinformativni apriorni porazdelitvi

Dobimo:

$$\mu \mid \sigma^2, x \sim N(\bar{x}, \sigma^2/n),$$

Aposteriorno porazdelitev σ^2 lahko zapisemo na tri ekvivalentne nacine, eden izmed njih:

$$(n-1)s^2/\sigma^2 \mid x \sim \chi^2_{n-1}.$$

Opazimo, da v aposteriorni porazdelitvi μ in σ^2 nista vec neodvisna, ceprav sta bila v apriorni.

Neinformativno porazdelitev si lahko predstavljamo nekako kot, da v konjugirani nastavimo κ_0 in ν_0 na 0.

Zgornji formuli sta dobro znani iz frekventisticne statistike.

3.4 Ali smo ocenili, kar nas je zanimalo?

Dejansko nas zanima robna (marginalna) aposteriorna porazdelitev μ.

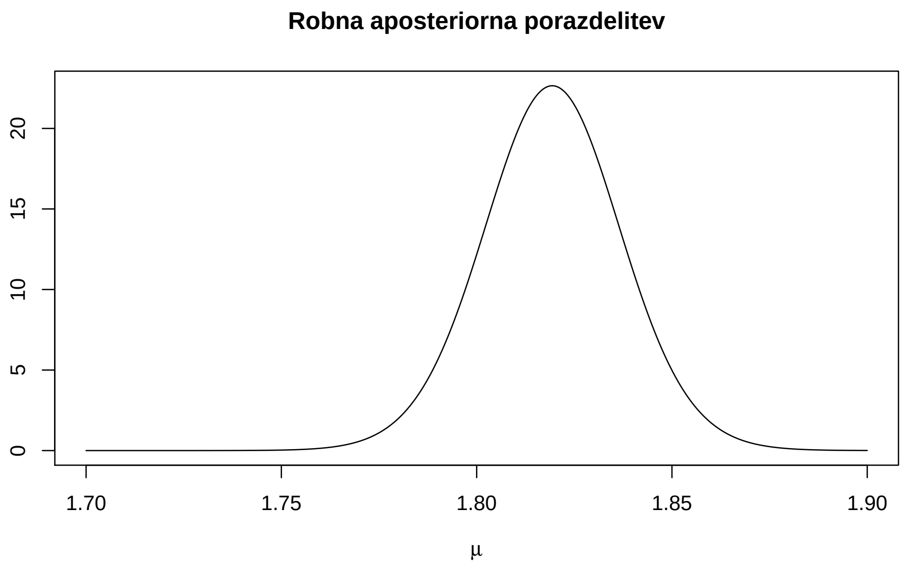
3.4.1 Robna aposterirona porazdelitev pri konjugirani apriorni porazdelitvi

V primeru konjugirane apriorne porazdelitve dobimo posploseno/nestandardizirano Studentovo porazdelitev:

$$\mu \mid x \sim t_{\nu_n}(\mu_n, \sigma_n^2/\kappa_n).$$

Narisemo v R-u:

```
mu <- seq(1.7, 1.9, 0.0001)
aposteriorna <- dt((mu-mu.n)/sqrt(sigma2.n/kappa.n), df = nu.n)/sqrt(sigma2.n/kappa.n)
plot(mu, aposteriorna, type="l", main="Robna aposteriorna porazdelitev",
     xlab = expression(mu), ylab = "")
```



3.4.2 Robna aposterirona porazdelitev pri neinformativni apriorni porazdelitvi

V primeru neinformativne apriorne porazdelitve dobimo posploseno/nestandardizirano Studentovo porazdelitev:

$$\mu \mid x \sim t_{n-1}(\bar{x}, s^2/n).$$

Oglejmo si, kaj velja za standardizirano odstopanje vzorcnega povprecja od populacijskega povprecja v naslednjih dveh pristopih:

- Bayesova statistika pri neinformativni apriorni porazdelitvi: $(\mu - \bar{x})/(s^2/n) \mid x \sim t_{n-1}$
- Frekventisticna statistika: $(\bar{x} - \mu)/(s^2/n) \mid \mu \sim t_{n-1}$

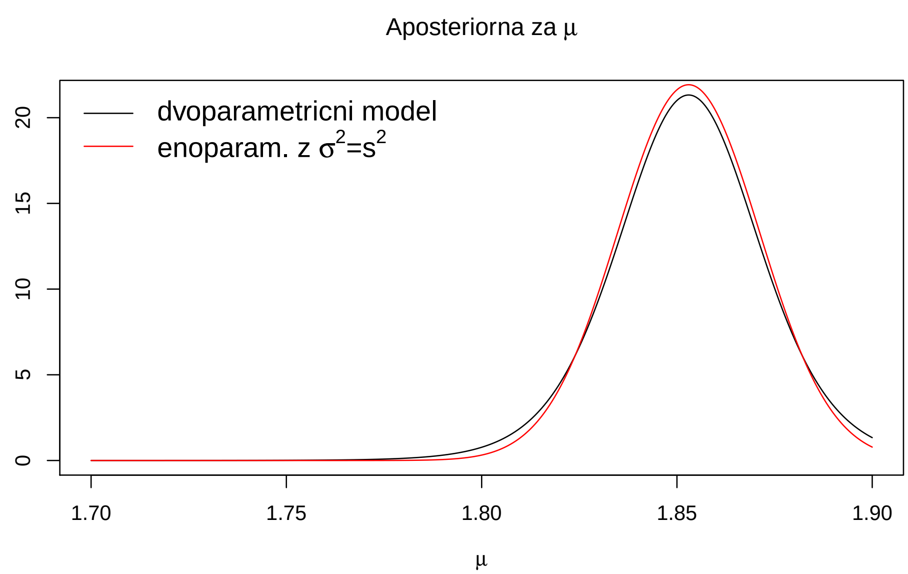
Negotovost o standardiziranemu odstopanju vzorcnega povprecja od populacijskega povprecja torej predstavlja enaka porazdelitev (tj. t_{n-1}) preden poznamo podatke (frekventisticna statistika) in po znanih podatkih (Bayesova statistika). Razlika pri tem je, da preden poznamo podatke dejansko ne poznamo niti $\bar{x}$ niti μ, po znanih podatkih pa poznamo $\bar{x}$, ki nam informira vedenje o μ.

Ce imamo v modelu $N(\mu, \sigma^2 = 0.1^2)$ neinformativno aposteriorno porazdelitev, potem je

$$\mu \mid x \sim N(\bar{x}, \sigma^2/n) = N(\bar{x}, 0.1^2/n).$$

Primerjamo rezultata iz obeh modelov (neinformativna apriorna v eno- ali dvoparametricnem modelu), kjer se zozimo na podvzorec prvih 10 visin (pri vecjem n razlika ne bi bila tako ocitna) in v modelu z znano varianco vzamemo kar $\sigma^2 = s^2$ (s cimer goljufamo, saj iz vzorca ocenimo varianco, ki naj bi bila znana vnaprej):

```
x.podvzorec <- x[1:10]

mu <- seq(1.7, 1.9, 0.0001)
aposteriorna.neinf <- dt((mu-mean(x.podvzorec))/sqrt(var(x.podvzorec)/n),
                         df = length(x.podvzorec)-1)/sqrt(var(x.podvzorec)/n)
aposteriorna.neinf.sigmaZnan <- dnorm(mu, mean = mean(x.podvzorec),
                                      sd = sqrt(var(x.podvzorec)/n))
plot(mu, aposteriorna.neinf, type="l",
     main=expression(paste("Aposteriorna za ",mu,sep="")),
     xlab = expression(mu), ylab = "")
lines(mu, aposteriorna.neinf.sigmaZnan, col="red")
legend("topleft", legend = c("dvoparametricni model",
       expression(paste("enoparam. z ",sigma^2,"=",s^2,sep=""))),
       col = c("black","red"), lty = 1, bty = "n", cex = 1.3)
```



Poanta: Ce goljufamo, potem je nasa ocena za μ manj variabilna – premalo variabilna!

3.5 Napovedovanje

Zanima nas, kaj lahko povemo o visini novega studenta ob upostevanju podatkov 30 studentov, tj. zanima nas **aposteriorna napovedna porazdelitev**.

V primeru konjugirane apriorne porazdelitve dobimo posploseno/nestandardizirano Studentovo porazdelitev:

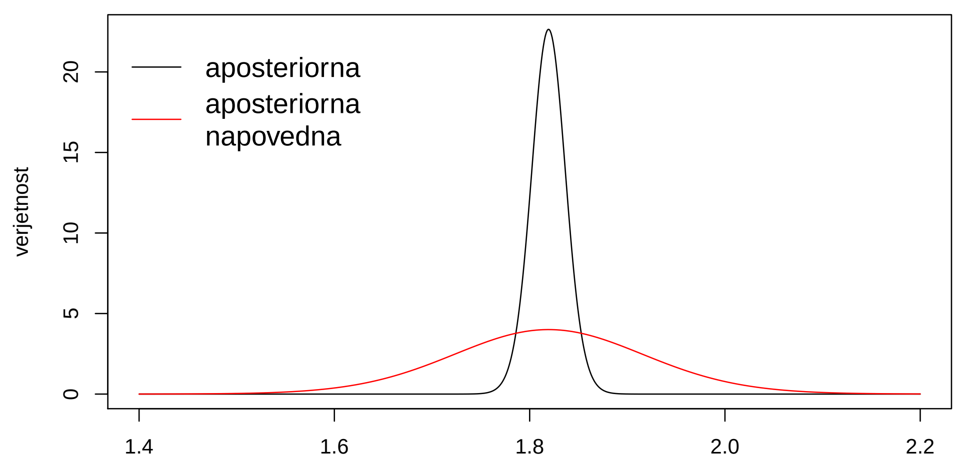
$$x_{nov} \mid x \sim t_{\nu_n}(\mu_n, \sigma_n^2/\kappa_n + \sigma_n^2).$$

V primeru neinformativne apriorne porazdelitve dobimo posploseno/nestandardizirano Studentovo porazdelitev:

$$x_{nov} \mid x \sim t_{n-1}(\bar{x}, s^2/n + s^2).$$

Poudarimo bistveno razliko med aposteriorno porazdelitvijo povprecne visine in aposteriorno napovedno porazdelitvijo za visino novega studenta (vzamemo primer s konjugirano apriorno porazdelitvijo):

```
mu <- seq(1.4, 2.2, 0.001)
aposteriorna <- dt((mu-mu.n)/sqrt(sigma2.n/kappa.n), df = nu.n)/sqrt(sigma2.n/kappa.n)
napovedna <- dt((mu-mu.n)/sqrt(sigma2.n/kappa.n+sigma2.n),
                df = nu.n)/sqrt(sigma2.n/kappa.n+sigma2.n)
plot(mu, aposteriorna, type="l",
     xlab = "", ylab = "verjetnost")
lines(mu, napovedna, col="red")
legend("topleft",  lty = 1,
       c("aposteriorna", "aposteriorna\nnapovedna"), col = c("black","red"),
       bty = "n", cex = 1.3)
```

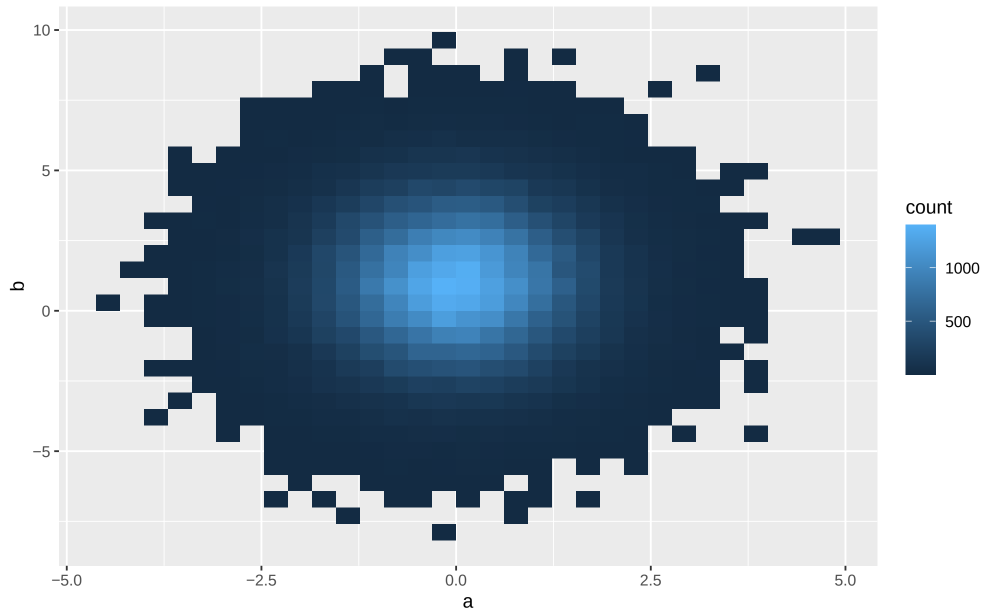


4 Naloga

Simulirajte iz aposteriorne porazdelitve, ki jo dobimo za nas primer pri zgoraj izbrani konjugirani apriorni porazdelitvi.

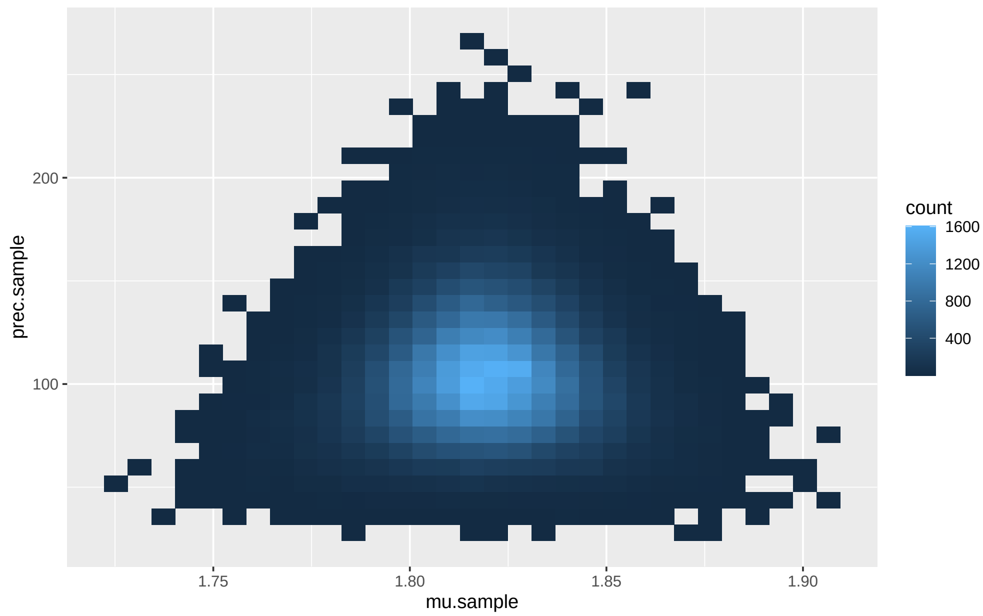
Za dobljeni vzorec narisite histrogram s knjiznico ggplot in ukazom stat_bin2d. Spodaj je primer za simulacijo iz bivariatne normalne porazdelitve, katere robni porazdelitvi sta neodvisni.

```
#install.packages("ggplot2")
library(ggplot2)
a <- rnorm(100000, 0, 1)
b <- rnorm(100000, 1, 2)
simulacija <- data.frame(a, b)
ggplot(simulacija, aes(a, b)) + stat_bin2d()
```



Resitev:

```
prec.sample <- rgamma(100000, nu.n/2, nu.n*sigma2.n/2)
mu.sample <- rnorm(100000, mu.n, sd = sqrt(1/(kappa.n*prec.sample)))

pod <- data.frame(mu.sample, prec.sample)
ggplot(pod, aes(mu.sample, prec.sample)) + stat_bin2d()
```




```
ggplot(pod, aes(mu.sample, 1/prec.sample)) + stat_bin2d()
```