EST–24107: Tarea 5

Carlos Lezama, Marco Medina,
Emiliano Ramírez y Santiago Villarreal

Miércoles, 17 de noviembre de 2021

Problema 1

Problema 2

Problema 3

Supongamos que queremos estimar $\int_0^1 e^{x^2} dx$.

Propongamos los siguientes dos estimadores:

1. $\hat{\theta}_1 = \frac{e^{u^2}\left(1+e^{1-2u}\right)}{2}$ y
2. $\hat{\theta}_2 = \frac{e^{u_1^2}+e^{u_2^2}}{2}$,

donde u, u_1 y u_2 son números aleatorios y $u_1 \perp u_2$.

Nótese que podemos reescribir

$$
\begin{aligned}
\hat{\theta}_1 &= \frac{e^{u^2}+e^{1-2u+u^2}}{2} \\
&= \frac{e^{u^2}+e^{(1-u)^2}}{2}
\end{aligned}
$$

Asimismo, podemos definir $h(x) = e^{x^2}$ tal que

$$\operatorname{Cov}\left(h(u), h(1-u)\right) < 0$$

como consecuencia de que $h(\cdot)$ es monótona y $\hat{\theta}_1$ se trata del promedio de dos variables antitéticas.

Para $\hat{\theta}_2$, es fácil ver que por independencia

$$\operatorname{Cov}\left(h(u_1), h(u_2)\right) = 0$$

y, por lo tanto, este estimador no alcanza reducción de varianza.

En consecuencia, el algoritmo definido por $\hat{\theta}_1$ es mejor que el descrito por $\hat{\theta}_2$.

De modo complementario, se puede observar una muestra numérica en la siguiente página.

```r
set.seed(1234)

theta1 <- function(n) {
  u <- runif(n)
  theta <- (exp(u ^ 2) + exp((1 - u) ^ 2)) / 2
  return(c(estimate = mean(theta),
           sd = sd(theta)))
}
theta2 <- function(n) {
  u1 <- runif(n)
  u2 <- runif(n)
  theta <- (exp(u1 ^ 2) + exp((u2) ^ 2)) / 2
  return(c(estimate = mean(theta),
           sd = sd(theta)))
}

theta1(1000000)
```

```
##  estimate        sd
## 1.4628581 0.1672392
```

```r
theta2(1000000)
```

```
##  estimate        sd
## 1.4628351 0.3356062
```

Problema 4

Problema 5

Problema 6

El número de reclamos que se harán en una aseguradora la próxima semana depende de un factor ambiental U. Si el valor de ese factor es $U = u$, entonces el número de reclamos tendrá distribución Poisson con media $\lambda = 15/(0.5 + u)$. Suponiendo que $U \sim \mathcal{U}(0, 1)$, definamos p como la probabilidad de que habrá al menos 20 reclamos la siguiente semana.

Para obtener una simulación cruda de p, podemos escribir

$$p = P(X > 20) = 1 - P(X \leq 20)$$

tal que

$$\begin{aligned}
P(X \leq 20) &= \int_0^1 P(X \leq 20 \mid u) du \\
&= \int_0^1 \sum_{i=0}^{20} \frac{e^{-\lambda}\lambda^i}{i!} du \\
&= \sum_{i=0}^{20} \frac{1}{i!} \int_0^1 e^{-\lambda}\lambda^i du \\
&= \sum_{i=0}^{20} \frac{1}{i!} F(i).
\end{aligned}$$

Entonces, podemos estimar cada integral $F(i)$, $\forall i = 0, 1, 2, \ldots, 20$ como sigue:

```r
f <- NULL

for (i in 0:20) {
  u <- runif(100000)
  y <- exp(-15 / (0.5 + u)) * (15 / (0.5 + u)) ^ i
  f[i + 1] <- mean(y)
}

p <- 1 - sum(f / factorial(0:20))

p

## [1] 0.2520797
```

Creamos la variable de control con la dependencia entre u y $X \mid u$ tal que

```
k <- 1000
n <- 100000

# Piloto
xu <- NULL
u <- runif(k)

for (i in 1:length(u)) {
  xu[i] <- rpois(1, 15 / (0.5 + u[i]))
}

a <- -lm(xu ~ u)$coeff[2]

# Simulación
u <- runif(n)
x <- NULL

for (i in 1:length(u)) {
  x[i] <- rpois(1, 15 / (0.5 + u[i]))
}

v <- x + a * (u - 0.5)
p <- mean(v > 20)

var(xu)

## [1] 42.44291

var(v)

## [1] 18.72857

p

## [1] 0.1956
```

Finalmente, para el caso de variables antitéticas:

```
f <- NULL

for (i in 0:20) {
  u <- runif(5000)
  u1 <- exp(-15 / (0.5 + u)) * (15 / (0.5 + u)) ^ i
  u2 <- exp(-15 / (0.5 + (1 - u))) * (15 / (0.5 + (1 - u))) ^ i
  f[i + 1] <- (mean(u1) + mean(u2)) / 2
}

p <- 1 - sum(f / factorial(0:20))

p
```

```
## [1] 0.2503782
```

Problema 7

Problema 8

Problema 9

Problema 10

Los siguientes datos corresponden a las horas adicionales de sueño de 10 pacientes tratados con un somnífero B comparado con un somnífero A:

```
sample <- c(1.2, 2.4, 1.3, 1.3, 0, 1, 1.8, 0.8, 4.6, 1.4)
```

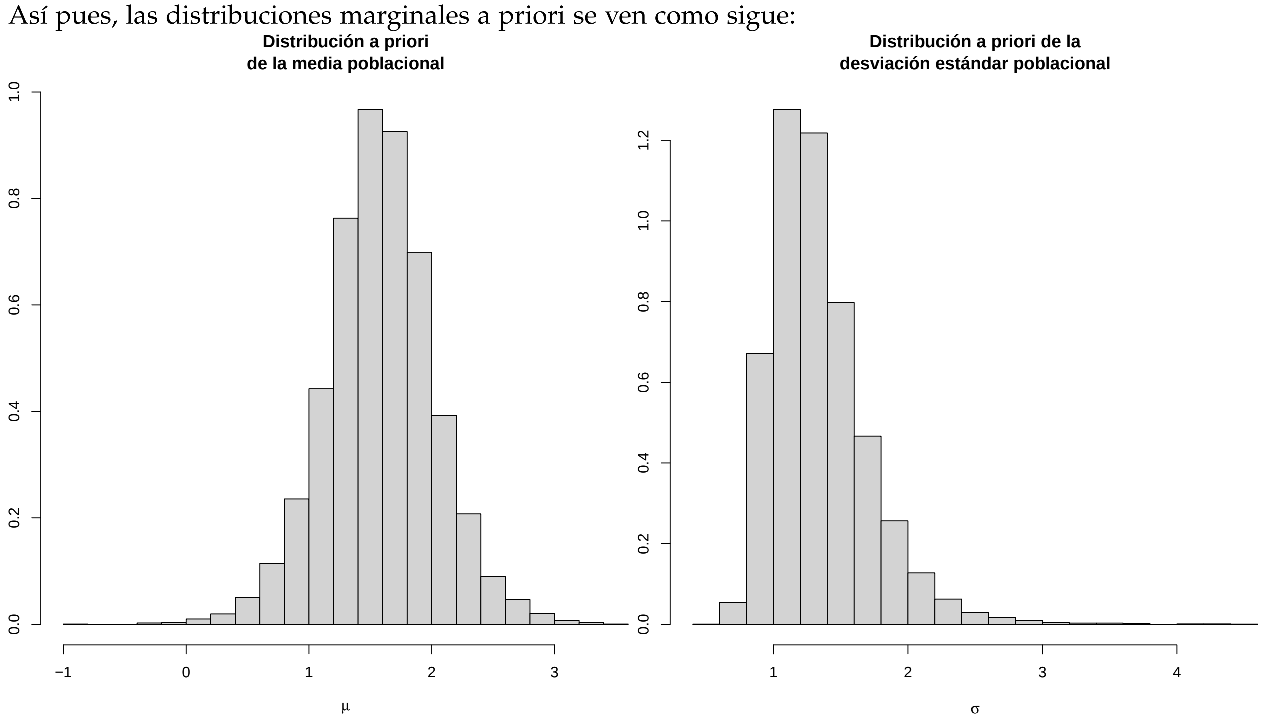
A priori, sabemos que $\bar{y} \sim \mathcal{N}\left(\mu, \sigma^2/n\right)$ y $(n-1)\dfrac{s_y^2}{\sigma^2} \sim \chi^2_{(n-1)}$.
De esta forma, podemos generar muestras aleatorias de

$$\sigma = \sqrt{(n-1)s_y^2/u_1}$$

tal que $u \sim \chi^2_{(n-1)}$. De forma análoga,

$$\mu = \bar{y} - \frac{u_2 \sigma}{\sqrt{n}}$$

donde $u_2 \sim \mathcal{N}(0,1)$.
Así pues, las distribuciones marginales a priori se ven como sigue:



Sea $m = \log(\sigma^2)$ y $p(m) \propto c$ tal que c es constante, por el método de la transformación inversa, $p(\sigma^2) \propto c/\sigma^2$ y $p(\sigma) \propto c/\sigma$. Si asumimos μ y σ independientes a priori tal que $p(\mu, \sigma) = p(\mu)p(\sigma)$, tenemos que $p(\mu, \sigma) \propto c/\sigma$.

Ahora bien, definiremos la siguiente función para obtener las curvas de nivel de las distribuciones posteriores conjuntas, así como sus distribuciones marginales al asumir que `sample` $\sim f_Y(y \mid \mu, \sigma)$.

```
posterior <- function(f, a, b, c, d, rate.1, rate.2) {
    aa <- seq(a, b, rate.1)
    bb <- seq(c, d, rate.2)
    post <- outer(aa, bb, f)
    rownames(post) = aa
    colnames(post) = bb
    post <- as.data.frame(post) %>%
        rownames_to_column(var = 'row') %>%
        gather(col, value,-row) %>%
        mutate(row = as.numeric(row),
               col = as.numeric(col))
    post <- post[!is.infinite(rowSums(post)), ]
    post <- na.omit(post)
    p <- ggplot(post, aes(
        x = row,
        y = col,
        z = value,
        fill = value
    )) +
        geom_tile() +
        geom_contour(color = 'black', size = 0.5) +
        scale_fill_viridis(option = 'mako',
                           direction = -1) +
        theme_minimal() +
        labs(x = expression(mu),
             y = expression(sigma),
             fill = NULL)
    p.mu <- ggplot(post, aes(x = row,
                             y = value)) +
        geom_point(size = 0.1) +
        theme_minimal() +
        labs(x = expression(mu),
             y = NULL) +
        theme(axis.text.y = element_blank())
    p.sigma <- ggplot(post, aes(x = col,
                                y = value)) +
        geom_point(size = 0.1) +
        theme_minimal() +
        labs(x = expression(sigma),
             y = NULL) +
        theme(axis.text.y = element_blank())
    return(list(p, p.mu, p.sigma))
}
```

Distribución normal

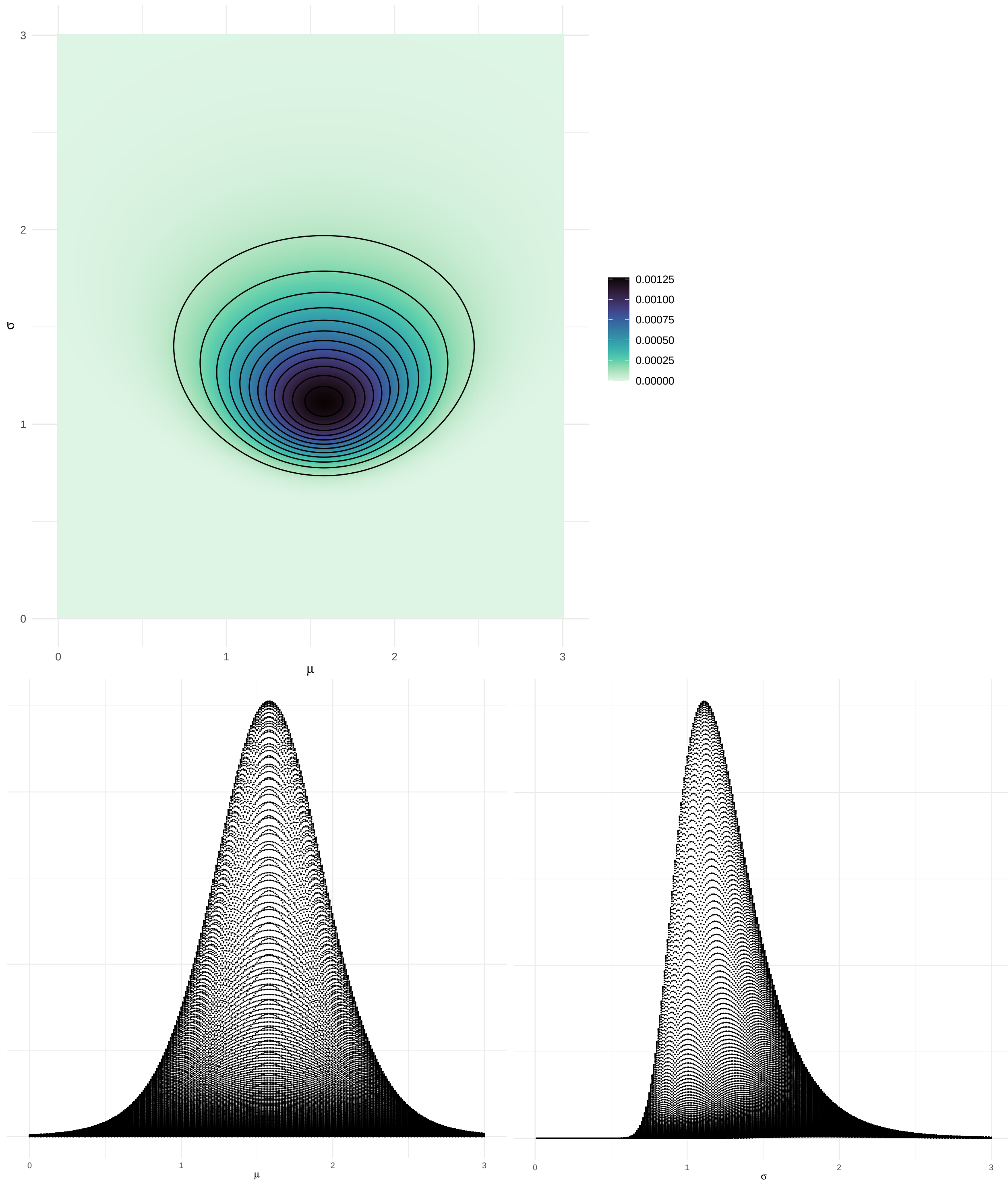
Sabemos $p(y \mid \mu, \sigma) \propto \ell(y \mid \mu, \sigma)$ donde $\ell(\cdot)$ es la función de verosimilitud correspondiente. Así pues,

$$
\begin{aligned}
p(\mu, \sigma \mid y) &\propto p(y \mid \mu, \sigma)\ p(\mu, \sigma) \\
&\propto \ell(y \mid \mu, \sigma)\ p(\mu, \sigma) \\
&= \frac{c}{\sigma} \left(2\pi\sigma^2\right)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \mu)^2\right) \\
&= c(2\pi)^{-n/2} \sigma^{-n-1} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \mu)^2\right) \\
&= \sigma^{-n-1} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \mu)^2\right) \\
&= \sigma^{-n-1} \exp\left(-\frac{1}{2\sigma^2} \left(\sum_{i=1}^{n} (y_i - \bar{y})^2 + n(\bar{y} - \mu)^2\right)\right) \\
&= \sigma^{-n-1} \exp\left(-\frac{1}{2\sigma^2} \left((n-1)s_y^2 + n(\bar{y} - \mu)^2\right)\right)
\end{aligned}
$$

De esta forma, podemos definir

```
bayes.norm <- function(a, b) {
  b ^ (-n - 1) *
    exp(-1 / (2 * b ^ 2) *
          (((n - 1) * (sample.sd ^ 2)) +
             (n * ((sample.mean - a) ^ 2
             ))))
}
```

Finalmente, podemos observar las siguientes curvas de nivel y distribuciones marginales $p(\mu \mid y)$ y $p(\sigma \mid y)$, respectivamente.


0.00125
0.00100
0.00075
0.00050
0.00025
0.00000
σ
μ
σ
μ

Adicionalmente, sabemos que

$$p(\mu \mid y) = \int_{\mathbb{R}_+} p(\mu, \sigma \mid y) d\sigma \implies \left. \frac{\sqrt{n}(\mu - \bar{y})}{s_y} \right| y \sim t_{(n-1)}$$

y

$$p(\sigma \mid y) = \int_{\mathbb{R}} p(\mu, \sigma \mid y) d\mu \implies \left. \frac{(n-1)s_y^2}{\sigma^2} \right| y \sim \chi^2_{(n-1)}$$

Distribución t

Sabemos que nuestra distribución t no centrada se define como sigue:

$$p(y \mid \hat{\mu}, \hat{\sigma}, v) = \frac{1}{\hat{\sigma}} k \left(1 + \frac{1}{v} \frac{(y-\hat{\mu})^2}{\hat{\sigma}^2}\right)^{-\frac{v+1}{2}}$$

donde

$$k = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{v\pi}\,\Gamma\left(\frac{v}{2}\right)},$$

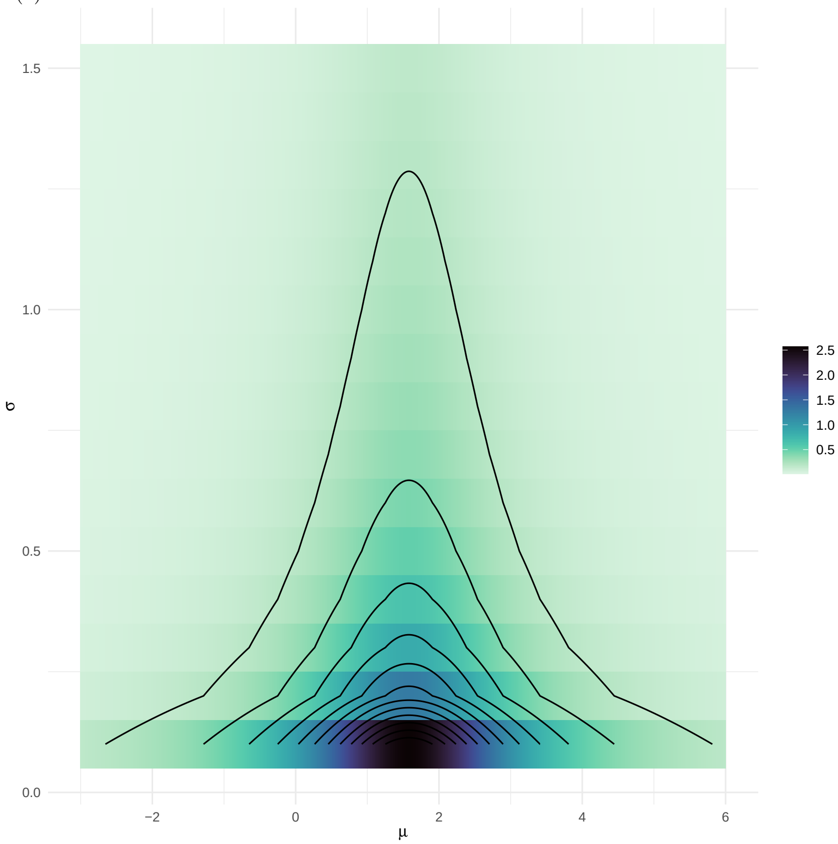
$\hat{\mu}$ es la media desconocida de una normal, $\hat{\sigma} = s_y/\sqrt{n}$ y v es fija.
Por lo tanto,

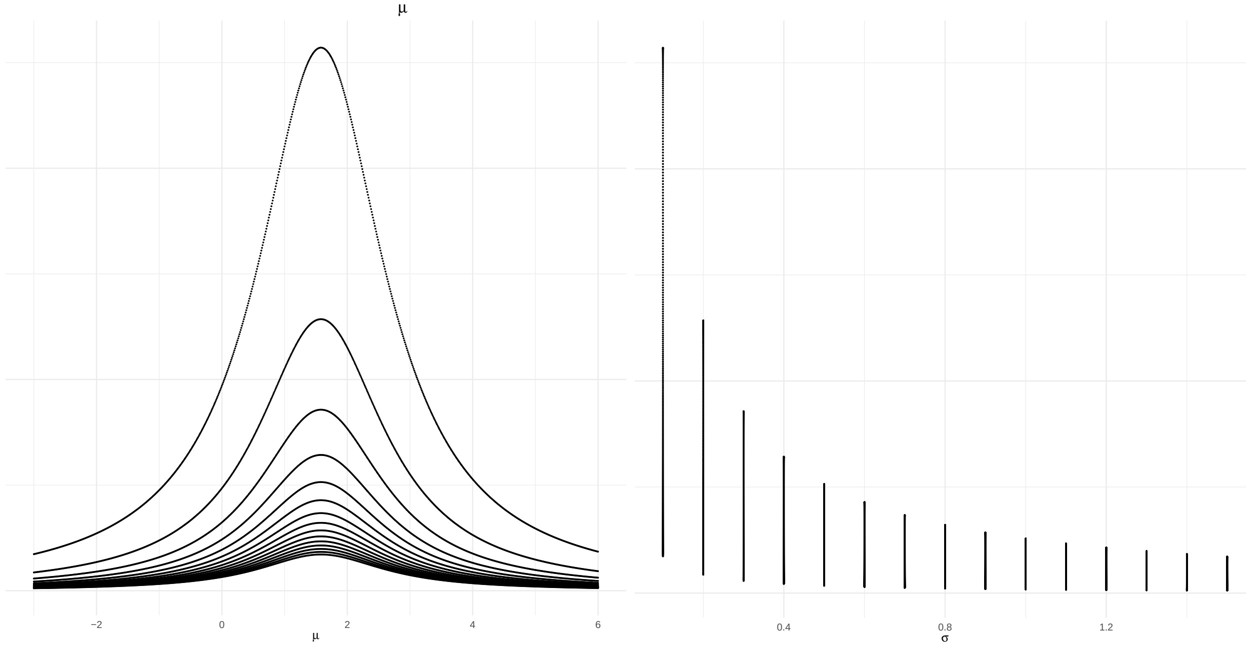
$$\begin{aligned}
p(\mu, \sigma \mid y) &\propto p(y \mid \mu, \sigma)\ p(\mu, \sigma) \\
&= k \frac{c}{\sigma} \frac{\sqrt{n}}{s_y} \left(1 + \frac{1}{v} \frac{n(y-\mu)^2}{s_y^2}\right)^{-\frac{v+1}{2}} \\
&= \frac{\sqrt{n}}{s_y \sigma} \left(1 + \frac{1}{v} \frac{n(y-\mu)^2}{s_y^2}\right)^{-\frac{v+1}{2}} \\
&= \frac{\sqrt{n}}{s_y \sigma} \left(1 + \frac{1}{v} \frac{\sum_{i=1}^{n}(y_i-\mu)^2}{s_y^2}\right)^{-\frac{v+1}{2}} \\
&= \frac{\sqrt{n}}{s_y \sigma} \left(1 + \frac{1}{v} \frac{(n-1)s_y^2 + n(\bar{y}-\mu)^2}{s_y^2}\right)^{-\frac{v+1}{2}}
\end{aligned}$$

Finalmente, en la siguiente página, definimos las funciones para 1 y 3 grados de libertad, respectivamente.

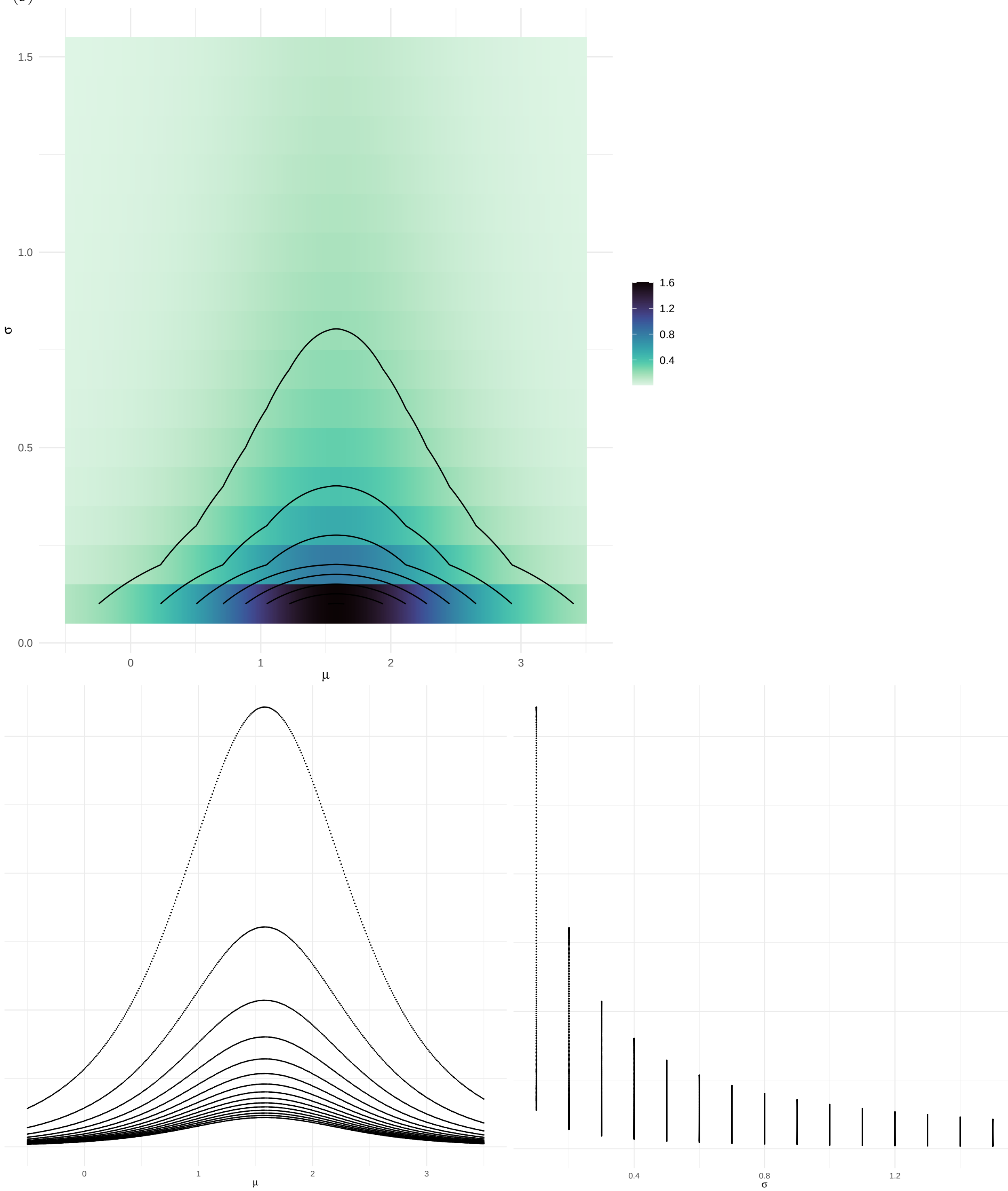
```
bayes.t.1 <- function(a, b) {
  (sqrt(n) / (b * sample.sd)) *
    (1 + ((((n - 1) * (sample.sd ^ 2)
    ) +
      (n * ((sample.mean - a) ^ 2
      ))) /
      (sample.sd ^ 2))) ^ (-1)
}

bayes.t.3 <- function(a, b) {
  (sqrt(n) / (b * sample.sd)) *
    (1 + (((((n - 1) * (sample.sd ^ 2)
    ) +
      (
        n * ((sample.mean - a) ^ 2)
      )) /
      (sample.sd ^ 2)) / 3)) ^ (-2)
}
```

$t_{(1)}$





$t_{(3)}$



Bernoulli

Sabemos,

$$
\begin{aligned}
p(y \mid \mu) &\propto \ell(y \mid \mu) \\
&= \prod_{i=1}^{n} \mu^{y_i}(1-\mu)^{1-y_i} \\
&= \mu^{\sum_{i=1}^{n} y_i}(1-\mu)^{\sum_{i=1}^{n}(1-y_i)} \\
&= \mu^{n\bar{y}}(1-\mu)^{n(1-\bar{y})}
\end{aligned}
$$

Además, $p(\mu) = c$. Por consiguiente,

$$
\begin{aligned}
p(\mu \mid y) &\propto p(y \mid \mu)p(\mu) \\
&\propto c \cdot \mu^{n\bar{y}}(1-\mu)^{n(1-\bar{y})} \\
&\propto \mu^{n\bar{y}}(1-\mu)^{n(1-\bar{y})}
\end{aligned}
$$

Finalmente, podemos observar su distribución posterior: