Tarea 1.2 Econometría

Wolfgang A. Resendiz Martínez

2023-02-20

Tarea 1

6. Suponga que el peso neto por lata en una marca de sopa tiene una $\mu = 565g$ y $\sigma^2 = 15g^2$. Suponiendo que la distribución de los pesos se asemeja a una distribución normal. Si se toma una muestra aleatoria de 9 latas y se registra el peso ¿cuál es la probabilidad de que la media muestral esté entre 560 y 568g? Provea un gráfico de la distribución normal en donde se señale con un área el segmento de la distribución que cubre el área pertinente a este ejercicio. Realícelo empleando el software R y la función `pnorm`.

```
library(tidyverse)
```

- Cálculo de la probabilidad de que la media muestral esté entre 560 y 568

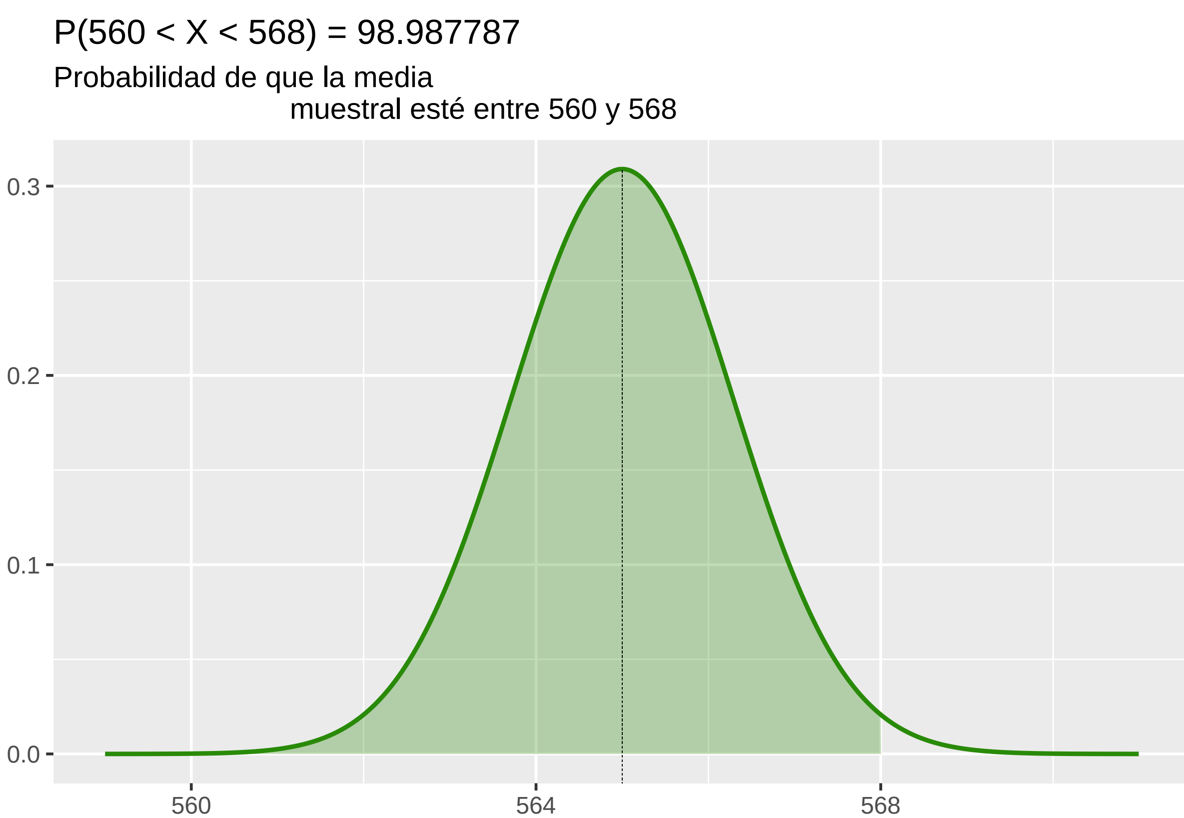
```
tamano_muestra = 9
media_pob = 565
var_pob = 15
desv_muestral = sqrt(var_pob/(tamano_muestra))

cota_inferior = 560
cota_superior = 568

p_intervalo = (pnorm(cota_superior,media_pob,desv_muestral)
               -pnorm(cota_inferior,media_pob,desv_muestral))*100

print(p_intervalo)
```

```
## [1] 98.98779
```

Grafica:

```
sigma_5 <- 5 * desv_muestral
cota_inf_sigma5 <- round((media_pob - sigma_5), 0)
cota_sup_sigma5 <- round((media_pob + sigma_5), 0)
step_intervalo_sigma5 <- (2 * sigma_5 / 1000)

x <- seq(from = cota_inf_sigma5, to = cota_sup_sigma5,
         by = step_intervalo_sigma5)
y <- dnorm(x, mean = media_pob, sd = desv_muestral)

datos_normal_1 <- data.frame(x, y)
datos_truncos <- datos_normal_1 %>% filter(x >= cota_inferior
                                           & x <= cota_superior)

graf_problem_6 <- datos_normal_1 %>% ggplot(aes(x = x, y = y)) +
  geom_line(
```

```
    data = datos_normal_1,
    aes(x = x, y = y),
    linewidth = 0.8,
    color = "#298A08"
  ) +
  geom_area(
    data=datos_truncos,
    aes(y = y),
    fill = "#298A08",
    alpha = 0.3
  ) +
  geom_segment(
    x = media_pob, xend = media_pob,
    y = -0.1, yend = dnorm(media_pob,media_pob,desv_muestral),
    linetype = "dashed",
    linewidth = 0.1,
  )+
labs(x = "", y = "") +
  ggtitle(label = sprintf("P(%d < X < %d) = %f" ,
                          cota_inferior, cota_superior, p_intervalo),
          subtitle = sprintf("Probabilidad de que la media
                             muestral esté entre %d y %d",
                             cota_inferior, cota_superior))

graf_problem_6
```



---

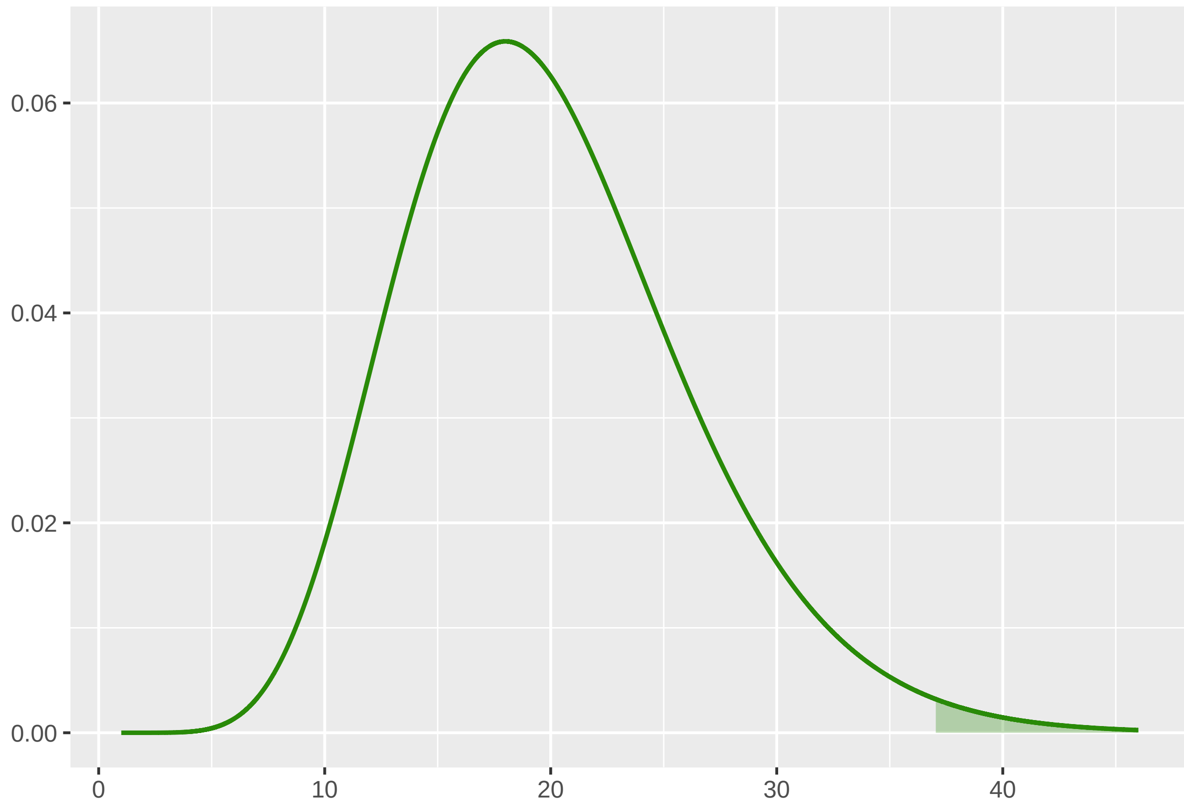
7. Un proceso industrial produce mosaicos de $10cm$ de ancho, con una desviación estándar (σ) de $0.9cm$.

Suponga que la muestra es aleatoria de una población normal. Si se toma una muestra de tamaño 21, calcule $P(S^2 > 1.5cm^2)$. Provea un gráfico de la distribución Chi-Cuadrada en donde se señale con un área el segmento de la distribución que cubre el área pertinente a este ejercicio. Realícelo empleando el software R y la función `pchisq`.

```
tamano_muestra_7 = 21
media_pob_7 = 10
desv_est_pob_7 = 0.9
var_pob_7 = desv_est_pob_7^2

cota_inferior_7 = 1.5
cota_superior_7 = NA

cota_inf_chisq = ((tamano_muestra_7-1)/(var_pob_7))*cota_inferior_7

print(cota_inf_chisq)
```

```
## [1] 37.03704
```

```
library(tidyverse)
cota_inf_sigma_7 <- round((media_pob_7 - 10 * desv_est_pob_7), 0)
cota_sup_sigma_7 <- round((media_pob_7 + 40 * desv_est_pob_7), 0)
step_intervalo_sigma5 <- (2 * desv_est_pob_7 / 1000)

x_7 <- seq(from = cota_inf_sigma_7, to = cota_sup_sigma_7,
           by = step_intervalo_sigma5)
y_7 <- dchisq(x_7, df = (tamano_muestra_7-1))

datos_normal_7 <- data.frame(x_7, y_7)
datos_truncos_7 <- datos_normal_7 %>% filter(x_7 >= cota_inf_chisq)

graf_problem_6 <- datos_normal_7 %>% ggplot(aes(x = x_7, y = y_7)) +
  geom_line(
    data = datos_normal_7,
    aes(x = x_7, y = y_7),
    linewidth = 0.8,
    color = "#298A08"
  ) +
  geom_area(
    data = datos_truncos_7,
    aes(y = y_7),
    fill = "#298A08",
    alpha = 0.3

  )+
labs(x = "", y = "")

graf_problem_6
```



---

8. **Considerando** el ejercicio 6 con $S^2 = 15$ y suponiendo que no se conoce el parámetro poblacional de σ^2.

   1. ¿Qué variable aleatoria conviene más para calcular la probabilidad del evento de interés?

   Una t de student

   2. Calcular la probabilidad del evento de interés. Provea gráficos que comparen las distribuciones de probabilidad involucradas.

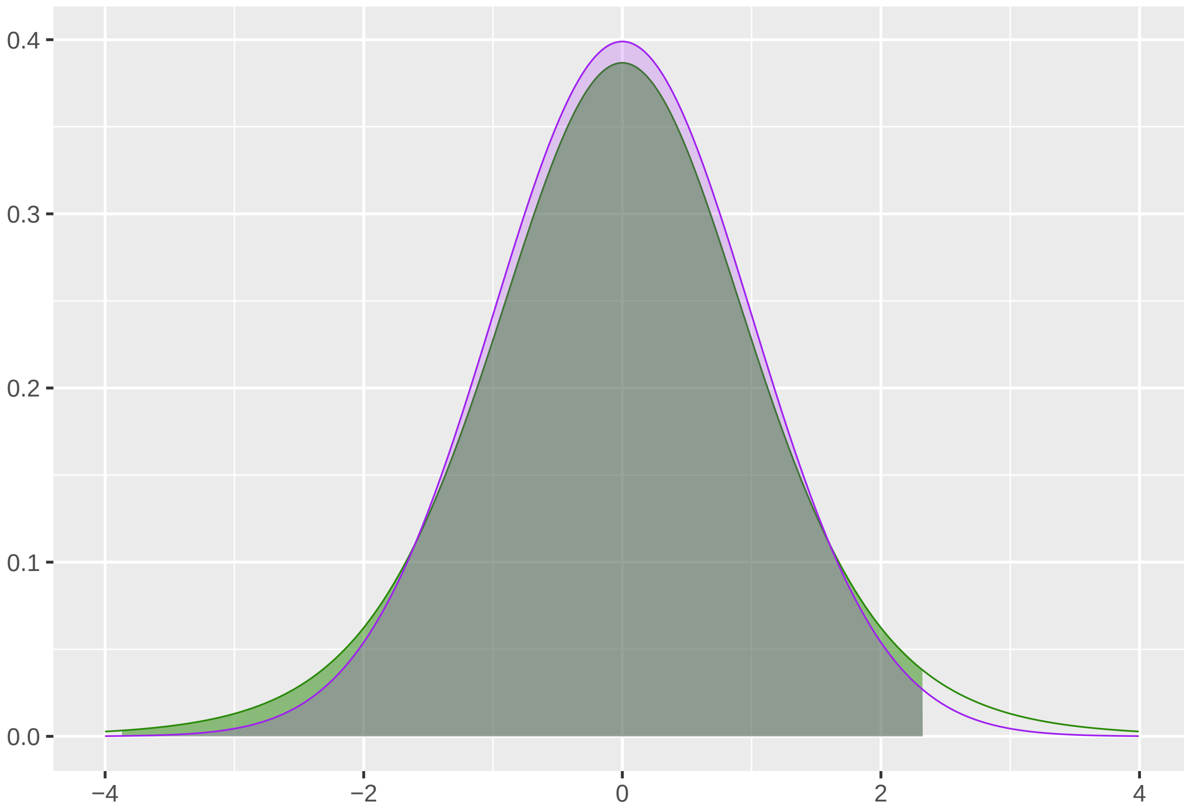
```
tamano_muestra_8 = 9
media_pob_8 = 565
var_mues_8 = 15

cota_inferior_8 = 560
cota_superior_8 = 568

cota_inf_t = ((cota_inferior_8 - media_pob_8)
              / sqrt(var_mues_8 / tamano_muestra_8))
cota_sup_t = ((cota_superior_8 - media_pob_8)
              / sqrt(var_mues_8 / tamano_muestra_8))


print(cota_inf_t)
```

```
## [1] -3.872983
```

```
print(cota_sup_t)
```

```
## [1] 2.32379
```

```r
library(tidyverse)

cota_inf_sigma_8 <- -4
cota_sup_sigma_8 <- 4
step_intervalo_sigma5 <- (2 * sqrt(var_mues_8) / 1000)

x_8 <- seq(from = cota_inf_sigma_8, to = cota_sup_sigma_8,
           by = step_intervalo_sigma5)
y_8 <- dt(x_8, df = (tamano_muestra_8 - 1))

y_6 <- dnorm(x_8, mean = 0, sd =  1)



datos_t_8 <- data.frame(x_8, y_8, y_6)
datos_truncos_t_8 <- datos_t_8 %>% filter(x_8 >= cota_inf_t & x_8 <= cota_sup_t)


graf_problem_8 <- datos_t_8 %>% ggplot(aes(x = x_8, y = y_8)) +
  geom_line(
    data = datos_t_8,
    aes(x = x_8, y = y_8),
    linewidth = 0.3,
    color = "#298A08"
  ) +
  geom_area(
    data = datos_truncos_t_8,
    aes(y = y_8),
    fill = "#298A08",
    alpha = 0.5

  )+
        geom_line(
    data = datos_t_8,
    aes(x = x_8, y = y_6),
    linewidth = 0.3,
    color = "purple"
  ) +
  geom_area(
    data = datos_truncos_t_8,
    aes(y = y_6),
    fill = "purple",
    alpha = 0.2
  ) +
labs(x = "", y = "")


graf_problem_8
```



La probabilidad bajo la aproximación de la distribución t es:

```
dt(cota_sup_t, df = (tamano_muestra_8 - 1)) - dt(cota_inf_t, df = (tamano_muestra_8 - 1))
```

```
## [1] 0.03462014
```

---

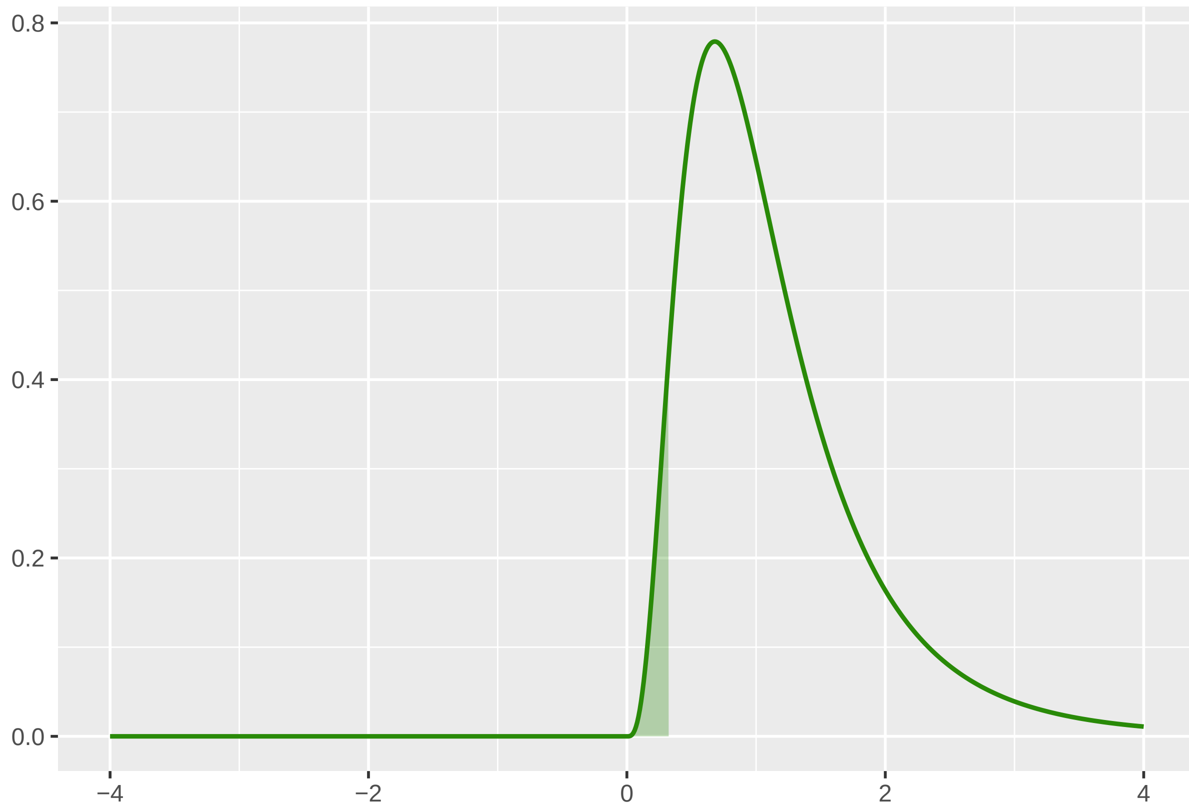
9. Un ingeniero agrónomo tiene dos variedades de maíz bajo investigación, p y r. Con base en el tiempo que el agrónomo tiene estudiando a ambas variedades asegura que la variedad p tiene mejor rendimiento promedio que la variedad r. Las variedades tienen sus pequeñas diferencias, pero no lo suficientemente grandes como para considerarlas diferentes. Si el agrónomo obtuviera una muestra aleatoria del rendimiento de 10 parcelas para la variedad p y una muestra aleatoria del rendimiento de 15 parcelas para la variedad r. ¿Cuál es la probabilidad de que la razón de varianzas $\frac{S_p^2}{S_r^2}$ sea menor a 0.5? Antes de contestar la pregunta anterior medite lo siguiente: bajo las consideraciones del agrónomo ¿cómo cree que será la probabilidad calculada de la pregunta anterior? Alta (probabilidad igual o mayor 0.8, mediana (probabilidad entre 0.3 y 0.8, sin tocar a estos) o pequeña (probabilidad entre 0 y 0.3).

```
tam_p = 10
tam_r = 15
cota = 0.5

cota_f = (tam_p - 1)/(tam_r - 1)*cota
cota_f
```

```
## [1] 0.3214286
```

```
cota_inf_sigma5 <- -4
cota_sup_sigma5 <- 4
step_intervalo_sigma5 <- (.01)
```

```r
x_9 <- seq(from = cota_inf_sigma5, to = cota_sup_sigma5,
         by = step_intervalo_sigma5)
y_9 <- df(x_9,tam_p-1, tam_r-1)

datos_9 <- data.frame(x_9, y_9)
datos_truncos_9 <- datos_9 %>% filter(x_9 <= cota_f)

graf_problem_6 <- datos_9 %>% ggplot(aes(x = x_9, y = y_9)) +
  geom_line(
    data = datos_9,
    aes(x = x_9, y = y_9),
    linewidth = 0.8,
    color = "#298A08"
  ) +
  geom_area(
    data = datos_truncos_9,
    aes(y = y_9),
    fill = "#298A08",
    alpha = 0.3
  ) +
  geom_segment(
    x = media_pob, xend = media_pob,
    y = -0.1, yend = dnorm(media_pob,media_pob,desv_muestral),
    linetype = "dashed",
    linewidth = 0.1,
  )+
labs(x = "", y = "")


graf_problem_6
```


0.8
0.6
0.4
0.2
0.0
−4
−2
0
2
4