Trabajo final - Métodos de Estimación 2023

Jesica Charaf e Ignacio Spiousas

2023-07-08

Presentación

La ditribución elegida para trabajar es la binomial negativa. Para el desarrollo utilizaremos la parametrización con r y μ que se muestra a continuación:

Sea $X \sim BN(r, \mu)$, entonces, su función de probabilidad puntual puede expresarse como:

$$Px(x) = \frac{\Gamma(x+r)}{x!\Gamma(r)} \left(\frac{r}{r+\mu}\right)^r \left(\frac{\mu}{r+\mu}\right)^x,$$

con $x = 1, 2, 3, ...$

Además, se tiene que: $\mathbb{E}(X) = \mu$ y $V(X) = \mu + \frac{\mu^2}{r}$.

Estimadores de momentos para r y μ

Obtención de los estimadores

Consideremos al conjunto de parámetros $\theta = (r, \mu)$ y que $X_1, ..., X_n \overset{iid}{\sim} BN(r, \mu)$. Como tenemos dos parámetros, para obtener los estimadores de momentos tendremos que utilizar el primer y segundo momento.

Primer momento

De plantear el primer momentos obtenemos:

$$\overline{X} = \mathbb{E}_{\hat{\theta}}(X)$$
$$\overline{X} = \hat{\mu}.$$

Segundo momento

Para el segundo momento tenemos que plantear:

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 = \mathbb{E}_{\hat{\theta}}(X^2),$$

Y para esto vamos a recordar que $V(X) = \mathbb{E}_{\hat{\theta}}(X^2) - (\mathbb{E}_{\hat{\theta}}(X))^2$, por lo tanto:

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 = \mathbb{E}_{\hat{\theta}}(X^2)$$
$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 = V(X) + (\mathbb{E}_{\hat{\theta}}(X))^2$$
$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 = \hat{\mu} + \frac{\hat{\mu}^2}{\hat{r}} + \hat{\mu}^2.$$

Resumiendo

Entonces, el sistema de ecuaciones que tenemos que resolver para hallar $\hat{r}$ y $\hat{\mu}$ es:

$$\overline{X} = \hat{\mu}$$
$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 = \hat{\mu} + \frac{\hat{\mu}^2}{\hat{r}} + \hat{\mu}^2.$$

El despeje

En el caso del estimador de $\hat{\mu}$, no hace falta despejar nado ya que queda directamente definido por el primer momento. Por otro lado, a $\hat{r}$ lo podemos despejar fácilmente del segundo momento reemplazando μ por $\overline{X}$:

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \hat{\mu} - \hat{\mu}^2 = \frac{\hat{\mu}^2}{\hat{r}}$$
$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X} - \overline{X}^2 = \frac{\overline{X}^2}{\hat{r}}$$
$$\hat{r} = \frac{\overline{X}^2}{\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X} - \overline{X}^2}$$

Entonces los estimadores de momentos son:

$$\hat{\mu}_{mo} = \overline{X}$$
$$\hat{r}_{mo} = \frac{\overline{X}^2}{\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X} - \overline{X}^2}$$

Consistencia

De nuevo, probar la consistencia del estimador de μ_{mo} es directo, ya que por LGN $\overline{X} \overset{c.s.}{\to} \mu$, entonces, $\hat{\mu}_{mo}$ es una estimador consistente para μ.

Para estudiar la consistencia de $\hat{r}_{mo}$ también vamos a utilizae LGN, pero vamos a tener que utilizar algunas propiedades. Por LGN sabemos que:

$$\sum_{i=1}^{n} X_i^2 \overset{c.s.}{\to} \mathbb{E}_{\hat{\theta}}(X^2) = \mu + \frac{\mu^2}{r} + \mu^2.$$

Además, como $\overline{X} \overset{c.s.}{\to} \mu$ y $g(t) = t^2$ es una función continua, podemos afirmar que $\overline{X}^2 \overset{c.s.}{\to} \mu^2$. Luego

$$\frac{\overline{X}^2}{\frac{1}{n}\sum_{i=1}^n X_i^2 - \overline{X} - \overline{X}^2} \xrightarrow{c.s.} \frac{\mu^2}{\mu + \frac{\mu^2}{r} + \mu^2 - \mu - \mu^2}$$
$$\frac{\overline{X}^2}{\frac{1}{n}\sum_{i=1}^n X_i^2 - \overline{X} - \overline{X}^2} \xrightarrow{c.s.} \frac{\mu^2}{\frac{\mu^2}{r}} = r.$$

Por lo tanto, $\hat{r}_{mo}$ es un estimador consistente para r.

Distribución asintótica

Para la distribución asintótica vamos a generar simulaciones de datos con distribución binomial negativa y parámetros poblacionales $r = 1$ y $\mu = 5$ y para cuatro tamaños de muestra, con $n = 10, 50, 100, 500$. Con cada simulación vamos a estimar los parámetros utilizando los estimadores de momentos $(\hat{\theta})$ y, a partir de ellos vamos a calcular los estimadores estandarizados de acuerdo a:

$$\frac{\hat{\theta} - \theta}{se(\hat{\theta})}.$$

Una vez obtenidos los valores estandarizados para $\hat{r}$ y $\hat{\mu}$ graficaremos los histogramas de ambas distribuciones comparados con la densidad de una distribución $\mathcal{N}(0, 1)$.

Primero defino las funciones que me generarán mis estimaciones de momentos

```
mu_mo <- function(x) {
  mean(x)
}

r_mo <- function(x) {
  (mean(x))^2/(mean(x^2) - mean(x) - mean(x)^2)
}
```

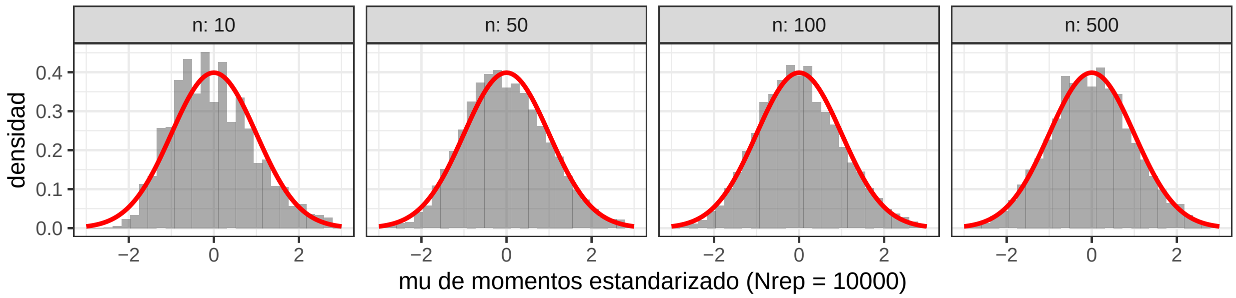
Luego, voy a realizar `Nrep` simulaciones del vector aleatorio X con $\{X_i, ..., X_n\} \sim BN(\mu = 5, r = 1)$:

```
ns <- c(1e1, 5e1, 1e2, 5e2)
Nrep <- 1e4
mu_pob <- 5
r_pob <- 1

est_rs <- vector(length = length(ns)*Nrep)
est_mus <- vector(length = length(ns)*Nrep)
est_ns <- vector(length = length(ns)*Nrep)

set.seed(12)
for (i in 1:length(ns)) {
  for (j in 1:Nrep) {
    data <- rnbinom(n = ns[i], size = r_pob, mu = mu_pob)
    est_ns[(i-1)*Nrep+ j] <- ns[i]
    est_mus[(i-1)*Nrep + j] <- mu_mo(data)
    est_rs[(i-1)*Nrep + j] <- r_mo(data)
  }
}
```

Luego estandarizamos los estimadores obtenidos con las muestra simuladas:

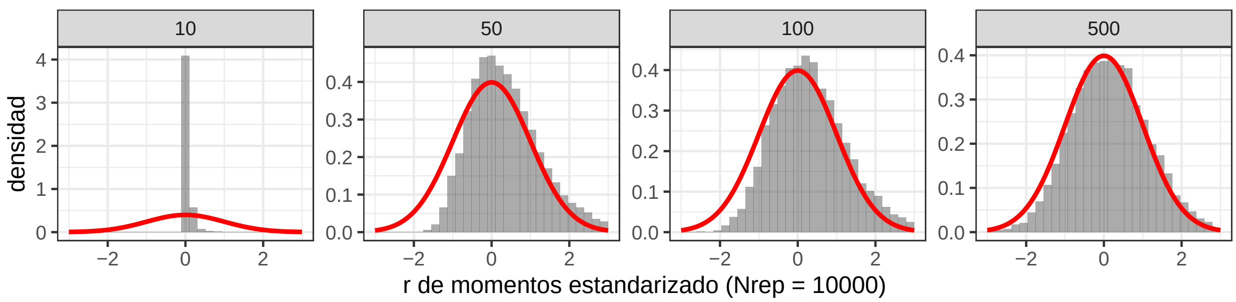
```
sim_mo <- tibble(n = est_ns,
                 mu = est_mus,
                 r = est_rs) %>%
  filter(r!=Inf) %>%
  group_by(n) %>%
  mutate(mu_est = (mu - mu_pob)/sd(mu),
         r_est = (r - r_pob)/sd(r))
```

Y, finalmente, vemos los histogramas de estos estimadores estandarizados comparados con una distribución $\mathcal{N}(0,1)$. Primero $\hat{\mu}_{mo,est}$:



En los histogramas puede verse que, aún para valores chicos de n, la distribución del estimador estandarizado $\hat{\mu}_{mo,est}$ es muy similar a la línea roja, es decir, su distribución asintótica es una $\mathcal{N}(0,1)$. Por otro lado, algo que podemos observar es que el estimador no pareciera tener un sesgo para ningún valor de n.

Y para $\hat{r}_{mo,est}$:



Y en este caso, y al contrario que para $\hat{\mu}_{mo,est}$, vemos que sólo para valores de n grandes la distribución de $\hat{r}_{mo,est}$ se aproxima a la $\mathcal{N}(0,1)$, mientras que para valores chicos de n muestra una distribución más concentrada y sesgada hacia la derecha.

Una cosa extra que podemos hacer, ya que tenemos simulaciones, es verificar la consistencia calculando el error cuadrático medio empírico (ECME):

```
sim_mo %>%
  group_by(n) %>%
  summarise(ECME_r = mean(r - r_pob)^2,
            ECME_mu = mean(mu - mu_pob)^2) %>%
  knitr::kable()
```

n	ECME_r	ECME_mu
10	0.2095478	7.51e-05
50	0.0244619	1.90e-06
100	0.0069814	1.77e-05

n	ECME_r	ECME_mu
500	0.0002816	1.90e-06

Puede verse que ambos estimadores son consistentes aunque, como se ve en los histogramas, $\hat{\mu}_{mo}$ parece ser un estimador insesgado mientras que $\hat{r}_{mo}$ parece ser un estimador asintóticamente insesgado, pero con un sesgo a derecha para valores chicos de n.

Estimadores de máxima verosimilitud para r y μ

Para hallar los estimadores de máxima versosimilitud utilizaremos el comando `fitdistr` del paquete {MASS}. Al igual que hicimos con los estimadores de momentos, vamos

Obtención de los estimadores

A modo de ejemplo obtendremos los parámetros $\hat{r}_{mv}$ y $\hat{\mu}_{mv}$ a partir de una muestra aleatoria de variables iid distribuidas como binomial negativa con parámetros $r = 1$ y $\mu = 10$.

```
mu_mo <- function(x) {
  fitdistr(x, densfun = "negative binomial")[2]
}

r_mo <- function(x) {
  fitdistr(x, densfun = "negative binomial")[1]
}
```

Puedo verirficar que las estimaciones me den bien:

Consistencia

Para la consistencia vamos a repetir el procedimiento utilizado para la distribución asintótica de los estimadores de momentos pero calcularemos el ECME. Primero hagamos la simulación:

```
ns <- c(1e1, 5e1, 1e2, 5e2)
Nrep <- 1e4
mu_pob <- 5
r_pob <- 1

est_rs <- vector(length = length(ns)*Nrep)
est_mus <- vector(length = length(ns)*Nrep)
est_ns <- vector(length = length(ns)*Nrep)

set.seed(12)
for (i in 1:length(ns)) {
  cat(paste("n =", ns[i], "\n"))
  for (j in 1:Nrep) {
    data <- rnbinom(n = ns[i], size = r_pob, mu = mu_pob)
    est_mv <- fitdistr(data, densfun = "negative binomial")
    est_ns[(i-1)*Nrep+ j] <- ns[i]
    est_mus[(i-1)*Nrep + j] <- est_mv$estimate[2]
    est_rs[(i-1)*Nrep + j] <- est_mv$estimate[1]
  }
}
```

```
## n = 10
```

```
## n = 50
## n = 100
## n = 500
```

Luego, estandarizamos los estimadores obtenidos con las muestra simuladas:

```
sim_mv <- tibble(n = est_ns,
                 mu = est_mus,
                 r = est_rs) %>%
  filter(r!=Inf) %>%
  group_by(n) %>%
  mutate(mu_est = (mu - mu_pob)/sd(mu),
         r_est = (r - r_pob)/sd(r))
```

Y, por último, calculamos el ECME:

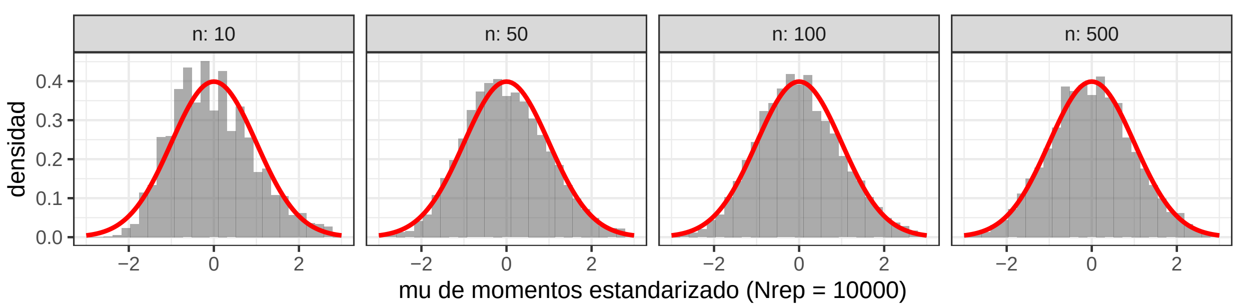
```
sim_mv %>%
  group_by(n) %>%
  summarise(ECME_r = mean(r - r_pob)^2,
            ECME_mu = mean(mu - mu_pob)^2) %>%
  knitr::kable()
```

n	ECME_r	ECME_mu
10	4.4644000	5.56e-05
50	0.0067779	1.90e-06
100	0.0018059	1.77e-05
500	0.0000514	1.80e-06

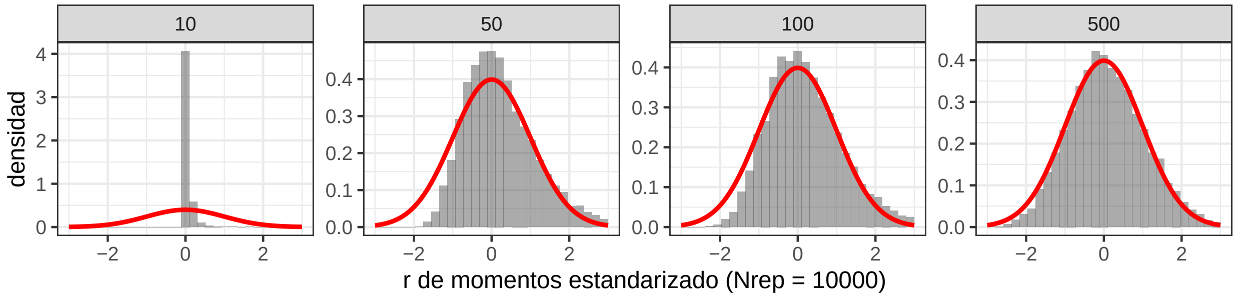
De la misma forma que en los estimadores de momentos, pareciera que, si bien ambos son consistentes, μ_{mv} es insesgado mientras que r_{mv} es asintóticamente insesgado.

Distribución asintótica

Después de estandarizar podemos ver los histogramas para $\hat{\mu}_{mv,est}$:



, y para $\hat{r}_{mv,est}$:



En los que, al igual que para los estimadores de momentos se ve que la distribución asintótica para ambos parámetros es una $\mathcal{N}(0,1)$, pero con diferencias para valores de n pequeños.

Aplicaciones y datos