Ejercicio: TCL

Leo Mansini 318/19

Clase 17 - 11 de junio

1. Calcular $E(\hat{p}_n)$ y $V(\hat{p}_n)$

$\hat{p}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ siendo X_i un experimento de Bernoulli de probababilidad:

$P(A) = 30/36 = 5/6 = 0.833$

$P(B) = 15/36 = 0.4167$

$P(C) = 1/6 = 0.167$

$E(\hat{p}_n)$ es una esperanza de una suma de variables aleatorias, por lo que es igual a la suma de las esperanzas de esas variables. X_i son i.i.d. para todo i entre 1 y n, entonces todas sus esperanzas son iguales:

$$E(\hat{p}_n) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n} nE(X_1) = E(X_1)$$

Y la esperanza de un experimento de Bernoulli de probabilidad p es igual a esa probabilidad.

Para $P(A)$ $E(\hat{p}_n) = 5/6 \approx 0.833$

Para $P(B)$ $E(\hat{p}_n) = 15/36 \approx 0.4167$

Para $P(C)$ $E(\hat{p}_n) = 1/6 \approx 0.167$

Análogamente busco la varianza.

$$V(\hat{p}_n) = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{1}{n^2} nV(X_1) = \frac{V(X_1)}{n}$$

Y la varianza de X_1 es $p(1-p)$

Para $P(A)$ $V(\hat{p}_n) = 5/36/1000 \approx 0.000139$

Para $P(B)$ $V(\hat{p}_n) = 55/432/1000 \approx 0.000127$

Para $P(C)$ $V(\hat{p}_n) = 5/36/1000 \approx 0.000139$

2. Estimar $E(\hat{p}_n)$ y $V(\hat{p}_n)$.

Puedo estimar la esperanza y la varianza usando los muestreos en pa.txt, pb.txt, y pc.txt.

Para estimar la esperanza, puedo hacer un promedio de las mediciones, es decir:

$$E(\tilde{p}_n) = \frac{\sum \hat{p}_n}{n}$$

Así,

Para pa.txt $E(\tilde{p}_n) =$

```
## [1] 0.8348163
```

Para pb.txt $E(\tilde{p}_n) =$

```
## [1] 0.419559
```

Para pc.txt $E(\tilde{p}_n) =$

```
## [1] 0.165342
```

Las tres estimaciones son buenas, distan en menos de 0.01 del valor real.

Para estimar la varianza uso su formula dependiente de la esperanza, y la aplico con la esperanza estimada:

$$V(\tilde{p}_n) = E(\tilde{p}_n^2) - E(\tilde{p}_n)^2$$

Entonces las varianzas estimadas para cada caso son:

Para pa.txt $V(\tilde{p}_n) =$

```
## [1] 0.0001820662
```

Para pb.txt $V(\tilde{p}_n) =$

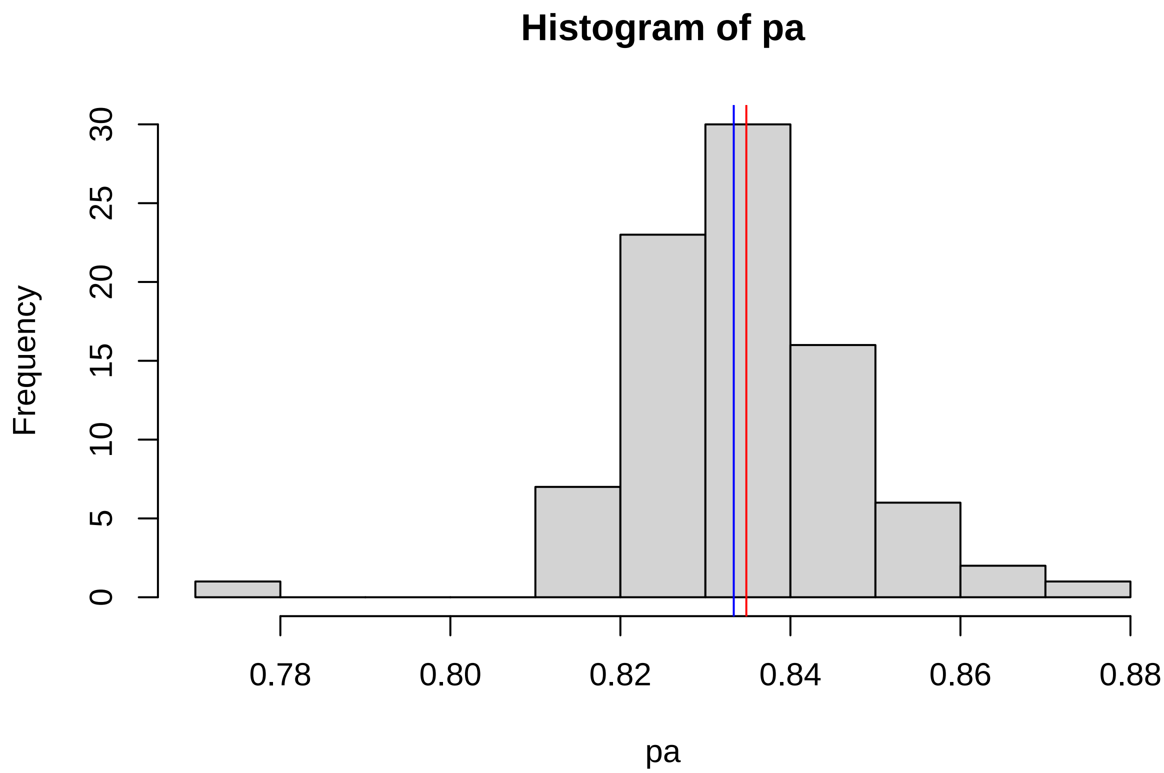
```
## [1] 0.0001852554
```

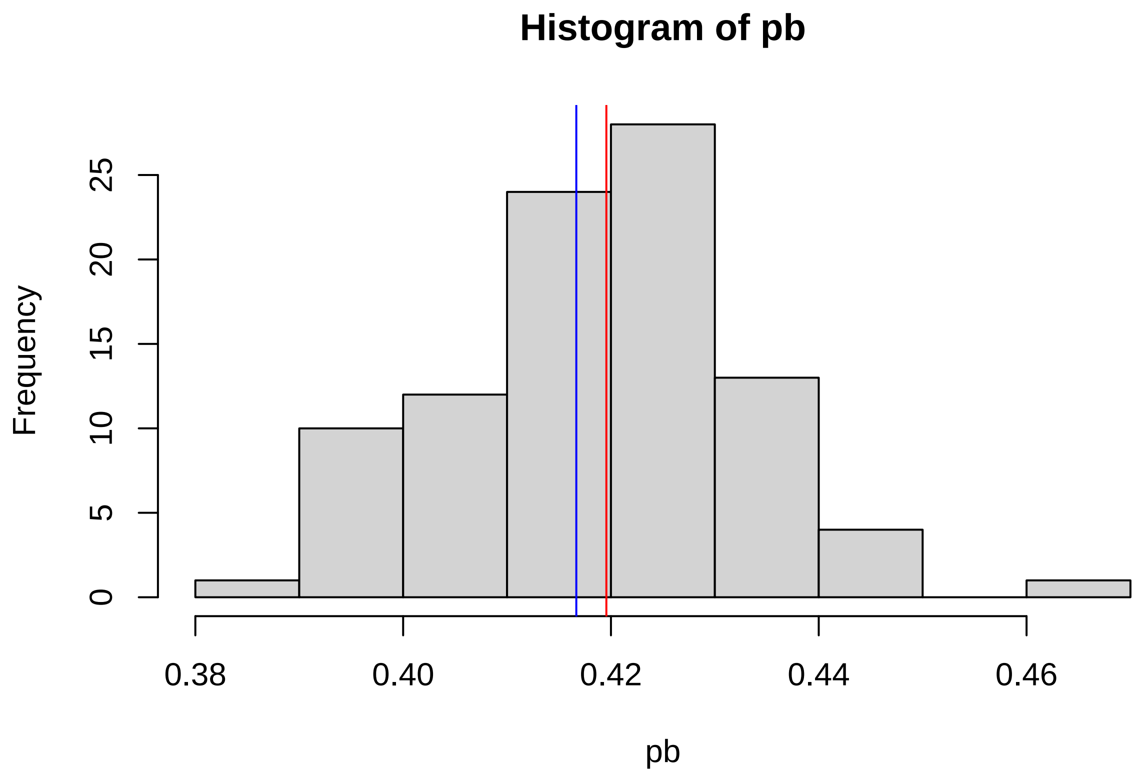
Para pc.txt $V(\tilde{p}_n) =$

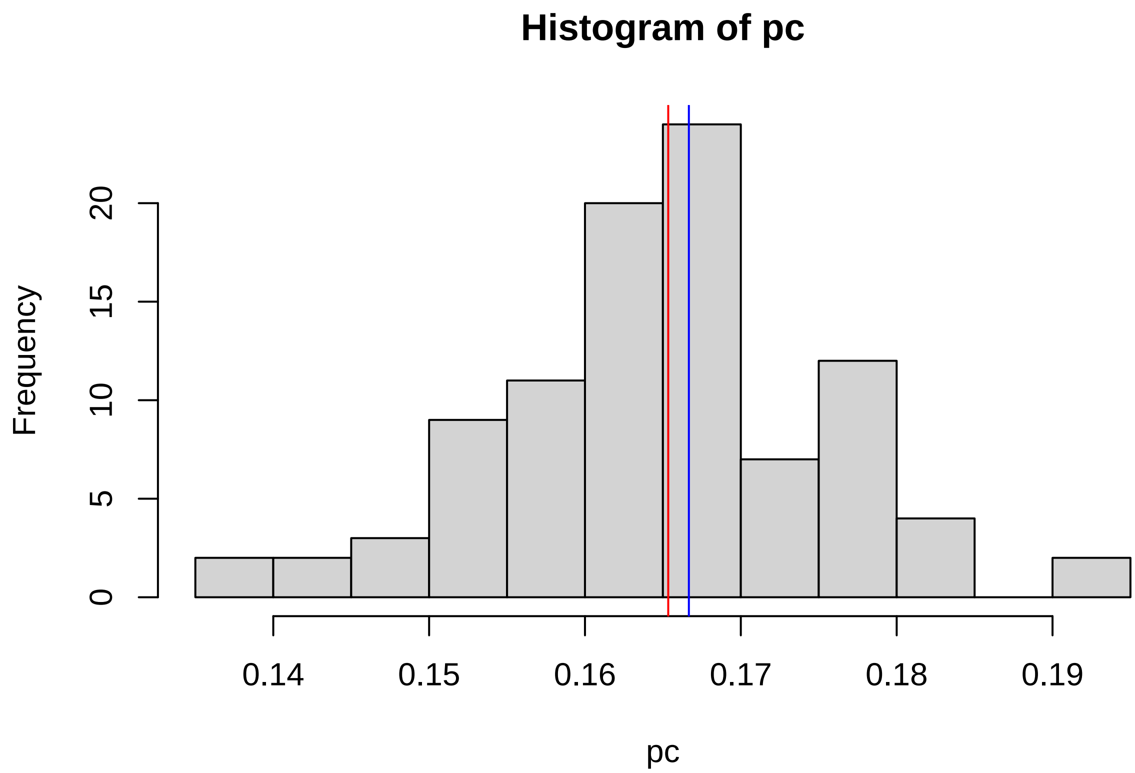
```
## [1] 0.0001106541
```

En este caso distan en un ~33%. Puede significar que la estimación de la varianza no converge a la real tan rapido como la esperanza, o que no son suficientes las muestras para aproximar.

3. Hago histogramas con las mediciones en pa.txt, pb.txt, pc.txt.


Histogram of pa
Frequency
0
5
10
15
20
25
30
0.78
0.80
0.82
0.84
0.86
0.88
pa


Histogram of pb
Frequency
0
5
10
15
20
25
0.38
0.40
0.42
0.44
0.46
pb

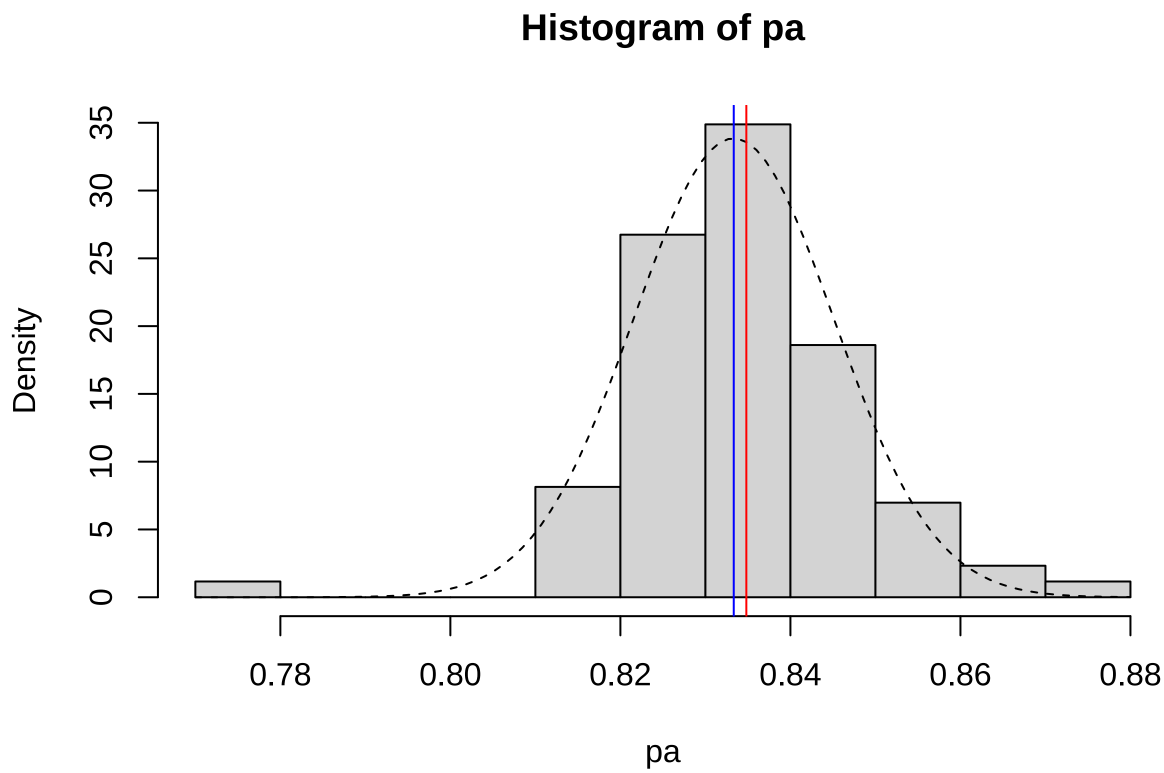


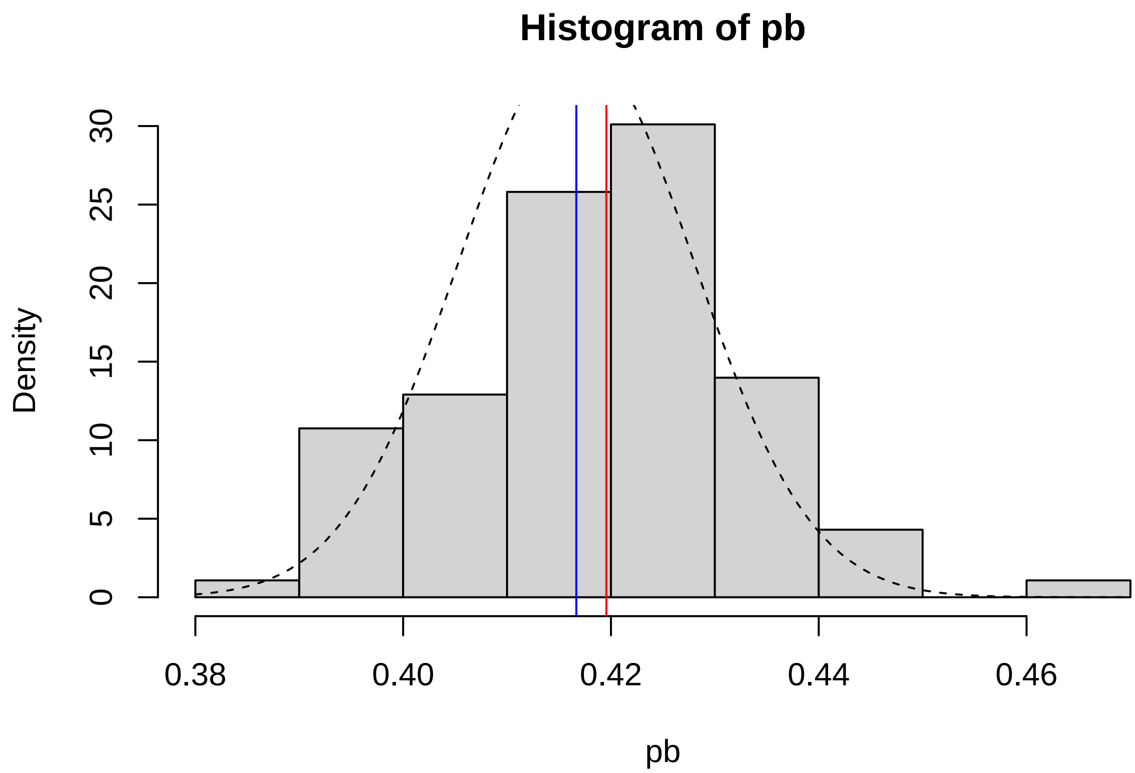
Se ve que el grafico tiene máximo cerca a la esperanza, estan marcadas las esperanzas reales en azul y las estimadas en rojo.

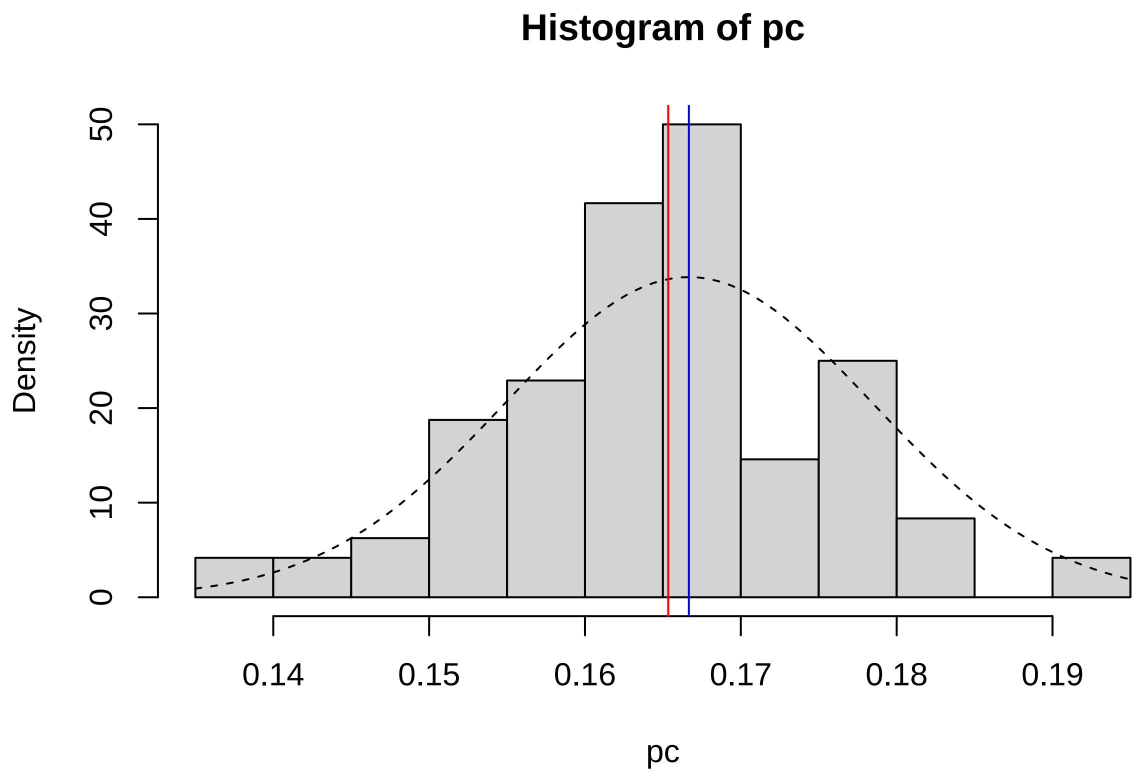
5. Agrego la densidad normal que aproxima a $\hat{p}_n$ sobre los histogramas hechos. Esta densidad es la de una variable Y de distribución normal de parametros $Y \sim \mathcal{N}(E(\hat{p}_n), \sqrt{V(\hat{p}_n)})$.

Es decir, la curva:

$$f(\hat{p}_n) = \frac{1}{2\pi\sqrt{V(\hat{p}_n)}} \exp\left(-\frac{1}{2}\left(\frac{\hat{p}_n - E(\hat{p}_n)}{\sqrt{V(\hat{p}_n)}}\right)^2\right)$$


Histogram of pa
Density
0
5
10
15
20
25
30
35
0.78
0.80
0.82
0.84
0.86
0.88
pa


Histogram of pb
Density
pb
0
5
10
15
20
25
30
0.38
0.40
0.42
0.44
0.46



En este caso, los gráficos estan normalizados, ya que la aproximación normal es una densidad.

Vemos que las curvas de densidad normal aproximan bien los histogramas de $P(A)$ y $P(B)$. En el caso de $P(C)$ la aproximación es menos acertada.

A medida que se aumenten la cantidad de muestras, el grafico tenderá a ser mas parecido a la curva de densidad normal.