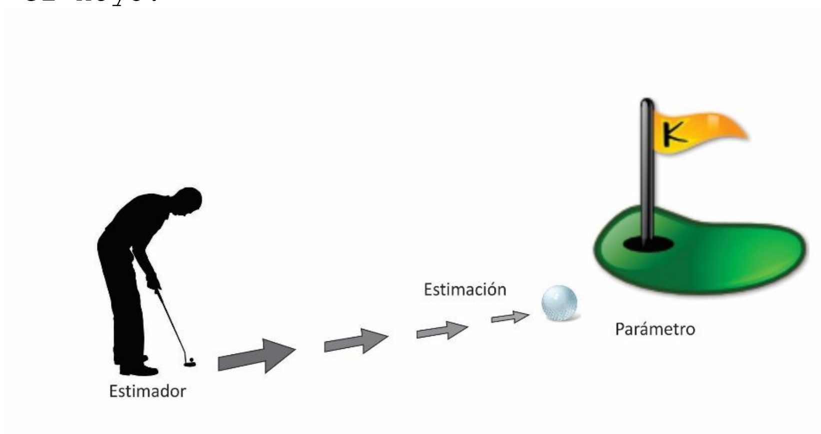
1.6 PROPIEDADES DE LOS BUENOS ESTIMADORES

Para introducir al tratamiento del tema de las propiedades de los buenos estimadores haremos una analogía con el golf. Nuestra intención es establecer el razonamiento empleado para determinar la bondad de un estimador.

La estimación es similar en muchos aspectos a golpear una pelotita hacia el hoyo. El estimador que genera las estimaciones es análogo al golpe, una estimación en particular es análoga al lugar donde se detiene la pelotita luego del golpe y el parámetro de interés es el hoyo.



Obtener una muestra de la población y estimar el valor del parámetro es equivalente a golpear la pelotita hacia el hoyo.

Supongamos que un jugador pega un golpe a la pelotita y emboca el hoyo. ¿Podemos concluir que éste es un jugador excelente? No, porque ninguno de nosotros estaría dispuesto a apostar que en un segundo disparo embocaría nuevamente el hoyo.

Pero, si una sucesión de 1,000,000 de tiradas dan en el hoyo, tendremos suficiente confianza en el jugador y apostaríamos una fuerte cantidad de dinero a favor del próximo acierto.

Con este simple ejemplo queremos aclarar que no se puede evaluar la bondad de un procedimiento de estimación sobre la base de un solo resultado.

Una medida de la bondad de la estimación sólo puede determinarse cuando el procedimiento se realiza muchas veces. Si esto ocurriera, veríamos cómo se distribuyen las llegadas de la pelotita alrededor del hoyo.

Por ejemplo, podríamos tener las siguientes situaciones:

Como las estimaciones son números, podemos evaluar la bondad del estimador construyendo una distribución de frecuencias de las estimaciones al repetirse el muestreo (cada golpe) y observando en qué forma se centra la distribución alrededor del parámetro de interés (el hoyo).

Así como un golpe de un jugador estará avalado por su performance a través de los años, en la práctica estadística, aunque trabajemos con una sola muestra y, en consecuencia, con un solo valor del estimador, la bondad del mismo estará respaldada por toda la teoría acerca de la distribución del estimador alrededor del parámetro que estima.

Con esta analogía en mente, pasaremos ahora a considerar las cuatro propiedades más importantes que deben cumplir los buenos estimadores.

Ellas son:

1) Insesgabilidad
2) Insesgabilidad de mínima varianza (eficiencia)
3) Consistencia
4) Distribución asintóticamente normal.

1) Insesgabilidad

La propiedad de insesgabilidad establece que el promedio o valor esperado de todos los valores posibles que puede tomar un estimador sea igual al parámetro estimado.

En otras palabras, si tomáramos todas las muestras posibles de un cierto tamaño de una población y calculáramos en cada muestra el estimador del parámetro de interés, podríamos obtener el promedio de estos valores. Si el valor encontrado coincide con el del parámetro, se dice que el estimador es insesgado.

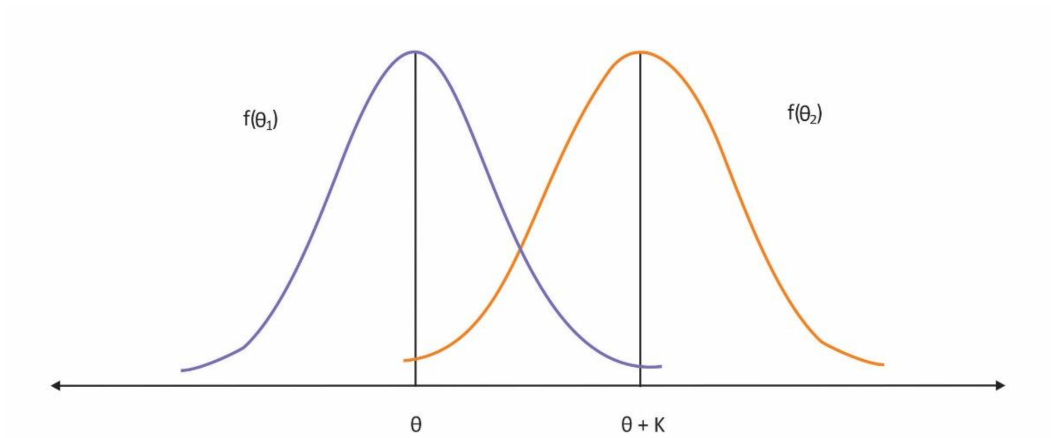
En símbolos:

θ parámetro
$\hat{\theta}$ estimador

Si $E(\hat{\theta}) = \theta$
$\hat{\theta}$ es un estimador "sin sesgo" o insesgado de θ.

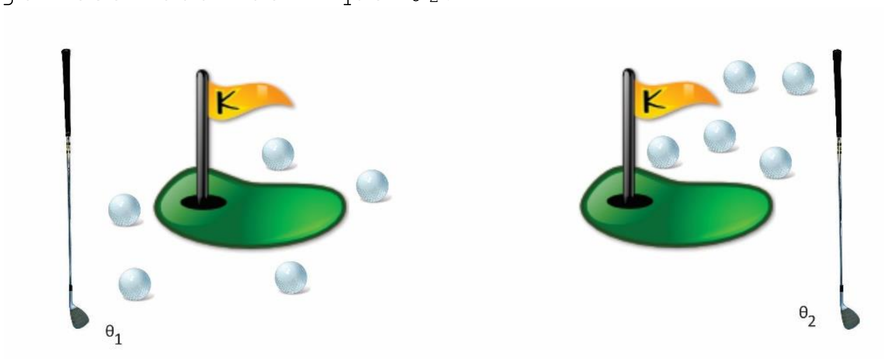
Si por el contrario: $E(\hat{\theta}) = \theta + K$
$\hat{\theta}$ es un estimador sesgado siendo K la magnitud del sesgo.

En general, si consideramos dos estimadores: $\hat{\theta}_1$ y $\hat{\theta}_2$ de un mismo parámetro θ y calculamos sus correspondientes distribuciones de frecuencias, podemos obtener una situación de este tipo:



En la figura anterior puede observarse que $\hat{\theta}_1$ es un estimador insesgado pues su esperanza coincide con el valor del parámetro θ, mientras que $\hat{\theta}_2$ es un estimador sesgado.

Por lo analizado hasta esta primera propiedad, podemos decir que $\hat{\theta}_1$ es mejor estimador de θ que $\hat{\theta}_2$.



En este último caso, el promedio está en otro lugar que no es el hoyo.

2) Insesgabilidad de varianza mínima

Esta propiedad dice que un buen estimador es aquel de mínima varianza con respecto a otros estimadores posibles del mismo parámetro.

En símbolos:

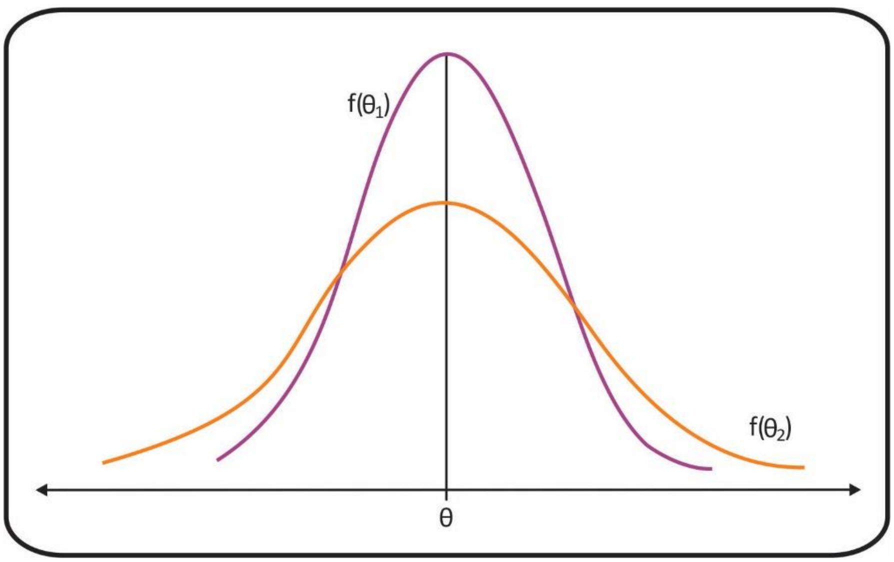
θ parámetro

$\hat{\theta}_1$ y $\hat{\theta}_2$ estimadores de θ tal que: $E(\hat{\theta}_1) = \theta$

$$E(\hat{\theta}_2) = \theta$$

Si $V(\hat{\theta}_1) < V(\hat{\theta}_2)$ entonces, $\hat{\theta}_1$ es mejor estimador de θ que $\hat{\theta}_2$.

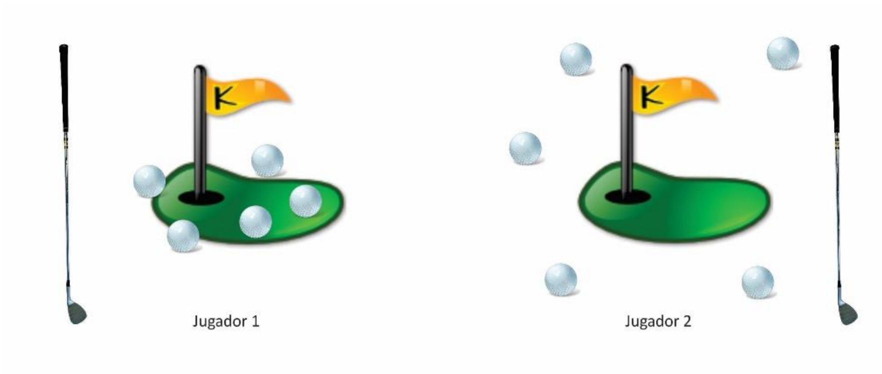
En general, si tenemos las siguientes distribuciones de frecuencias correspondientes a dos estimadores $\hat{\theta}_1$ y $\hat{\theta}_2$ de un cierto parámetro θ:



los dos estimadores son insesgados pues su esperanza coincide con el valor del parámetro θ. Pero, ¿qué podemos decir con respecto a su variabilidad? Evidentemente, los valores que toma $\hat{\theta}_1$ están mucho más concentrados alrededor del parámetro a estimar que los valores de $\hat{\theta}_2$.

Intuitivamente, un investigador confiará más en un estimador muestral que tenga menor posibilidad de alejarse mucho del verdadero valor del parámetro poblacional.

En el ejemplo de la pelotita, dos jugadores podrían haber realizado las siguientes jugadas sucesivas:



Evidentemente, nadie dudaría en apostar por el jugador 1 en sucesivos golpes, ha demostrado una mayor precisión para acercarse al hoyo.

Generalmente, suele decirse que un estimador de mínima varianza es un estimador eficiente.

3) Consistencia (Simple y de error cuadrático medio)

Intuitivamente un "buen" estimador sería aquel para el cual el riesgo disminuye cuando crece el tamaño de la muestra. Es decir, el estimador es mejor cuando se basa en una muestra de veinte observaciones, por ejemplo, que si se basa en una de sólo dos. Para ser más explícitos, supongamos que $\hat{\theta}_1$ es un estimador de θ basado en una muestra de tamaño 1, $\hat{\theta}_2$ un estimador de θ que se basa en una muestra de tamaño 2 y en general $\hat{\theta}_n$ un estimador de θ basado en una muestra de tamaño n.

La condición de consistencia (conocida como consistencia simple) establece que para muestras grandes $\hat{\theta}_n$ tiende a aproximarse a θ.

Existe no obstante otra condición de consistencia conocida como consistencia en error cuadrático que implica que tanto el sesgo como la varianza del estimador $\hat{\theta}_n$ tienden a cero a medida que aumenta el tamaño de la muestra.

4) Distribución asintóticamente normal

Se dice que un estimador es asintóticamente normal si además de ser insesgado y eficiente, cumple con la propiedad de tener distribución normal cuando el tamaño de la muestra se incrementa.

Por último, podemos decir que las propiedades 1 y 2 son satisfechas por estimadores calculados en muestras de cualquier tamaño, mientras que las propiedades 3 y 4 se cumplen solamente cuando los estimadores se calculan en base a observaciones provenientes de muestras grandes. Desarrollaremos ahora un ejemplo mediante el que podremos ver cómo el estimador media muestral cumple con todas las propiedades de un buen estimador.

Consideremos una población compuesta por 5 escuelas rurales en las que se ha registrado el número de maestros, obteniendo:

2, 3, 6, 8, 11

Si calculamos la media aritmética y la varianza correspondiente a la variable

X = cantidad de maestros por escuela

obtendremos los parámetros que caracterizan a esta distribución.

Estos resultados nos indican que el promedio de maestros en escuelas rurales en la población es de 6 maestros por escuela con una dispersión de 3.29 maestros por escuela.

Supongamos, ahora, que seleccionamos todas las muestras posibles de tamaño $n = 2$ por medio de un muestreo con reemplazo.

No.	Muestra 1	Muestra 2
1	2	2
2	2	3
3	2	4
4	2	6
5	2	8
6	2	11
7	3	2
8	3	3
9	3	4
10	3	6
11	3	8
12	3	11
13	4	2
14	4	3
15	4	4
16	4	6
17	4	8
18	4	11
19	6	2
20	6	3
21	6	4
22	6	6
23	6	8
24	6	11
25	8	2

26	8	3
27	8	4
28	8	6
29	8	8
30	8	11
31	11	2
32	11	3
33	11	4
34	11	6
35	11	8
36	11	11

El promedio de las muestras es:

$$\bar{x} = 5.66$$

Y la desviación estándar es:

$$s = 3.11$$

En el caso de que aumentemos el tamaño de la muestra, seguramente el promedio se irá acercando al parámetro poblacional. Mientras que la desviación estándar se reducirá, prueba de ello, es el valor obtenido en la muestra de las 5 escuelas comparado con las 36 muestras de tamaño 2.

1.7 PRECISIÓN DE LOS ESTIMADORES

Las estimaciones realizadas en forma puntual, es decir, calculando para cada caso particular el valor proveniente de una muestra, no lleva asociada alguna idea acerca del grado de aproximación que puede existir entre el valor del estimador y el valor del parámetro que se está estimando. Es decir, que no es posible conocer algo acerca del error que puede cometerse al afirmar que el parámetro desconocido se estima con esa función de observaciones muestrales definidas por cada estimador, ni de la confianza que puede depositarse en la sospecha de que el valor estimado se encuentre relativamente cerca del parámetro desconocido.

Precisamente, la diferencia existente entre el valor del estimador en una muestra particular y el verdadero valor del parámetro desconocido se llama “error de muestreo”.

$$\left|\hat{\theta} - \theta\right| \leq e$$

Este error es aleatorio porque depende del estimador que es una variable aleatoria. Esto significa que obviamente no es posible saber con exactitud en cuanto nos "equivocaremos" al estimar el parámetro a partir de una muestra.

Entonces, es interesante tener en cuenta el conocimiento de las distribuciones de probabilidad de algunos estimadores, es decir, conocer las funciones de los estimadores y de los parámetros, lo que permitirá analizar la posibilidad de establecer un intervalo aleatorio para estimar dónde se encuentra el valor del parámetro.

$$\left(\hat{\theta}-e,\hat{\theta}+e\right)$$

Cuya amplitud "*d*" es igual al doble del error máximo de estimación y al cual pueda asociarse una elevada probabilidad de que el parámetro sea interior al intervalo.

Si bien se ha perdido precisión en la estimación al referirse la misma a un intervalo y no a un valor puntual, se ha ganado en el conocimiento del error que puede cometerse al realizarla y del grado de confianza que la afirmación sea verdadera.