# Inferencia Estadística

- Inferencia Estadística
  - Estimación puntual
  - Estimación por intervalos
  - Estimación bayesiana
  - Contraste de hipótesis

# Inferencia Estadística: Estimación puntual

## Estimación puntual

- Introducción a la inferencia estadística
- Métodos de muestreo
- La estimación puntual
- Distribución de un estimador en el muestreo
- Propiedades de los estimadores
- Estimadores de máxima verosimilitud

# Introducción a la inferencia estadística

- La creación de modelos probabilísticos es un caso típico de razonamiento deductivo donde se generan las hipótesis generales sobre el mecanismo que origina los datos, generando así las distribuciones de probabilidades que originan los datos:
  - Por ejemplo definimos e un proceso de Bernoulli la variable binomial como: $y =$ número de elementos defectuosos al observar n observaciones. Ahora suponemos que hemos realizado n observaciones de las cuales r son defectuosas y n-r son aceptables (da igual el orden por la hipótesis de independencia). Con estas hipótesis dedujimos en capítulos anteriores la distribución de probabilidad binomial (razonamiento deductivo), y así con resto de distribuciones estudiadas anteriormente.
- El procedimiento inverso se realiza mediante la inferencia estadística, i.e. mediante las frecuencias observadas de una variable, extraer o inferir el modelo probabilístico que han generado los datos mostrando esas frecuencias (razonamiento inductivo).

# Introducción a la inferencia estadística

- Existen muchos tipos de inferencia estadística:
  - Según el objetivo del estudio: muestreo frente a diseño.
    - **Describir** variables y sus relaciones entonces se utilizan técnicas de muestreo.
    - **Contrastar** relaciones entre variables y **predecir** valores futuros, para ello se utilizan técnicas de diseño experimental (se fijan valores de cierta variables y se miden la respuesta que inducen otras).

# Introducción a la inferencia estadística

- Existen muchos tipos de inferencia estadística:
  - Por el método utilizado: métodos paramétricos v.s. no paramétricos.
    - **Paramétrico**: se supone que los datos provienen de una cierta distribución y se tienen muestras para estimar los parámetros de la misma.
    - **No paramétrico**: supones aspectos generales de la distribución (continua simétrica, etc.) y tratan de estimar o contrastar su estructura. Generalmente se estiman su forma mediante el suavizado los histogramas de los datos muestrales.

# Introducción a la inferencia estadística

- Existen muchos tipos de inferencia estadística:
  - Por la información considerada: enfoque clásico v.s. bayesiano.
    - **Clásico**: los parámetros son cantidades fijas desconocidas (sin información sobre ellos), y la inferencia utiliza solo la información de los datos muestrales.
    - **Bayesiano**: considera los parámetros como variables aleatorias y permite introducir información adicional sobre los mismos a través de una probabilidad a priori.

# Métodos de muestreo: muestra y población

- **Población**: conjunto homogéneo de elementos en los que se estudia una característica dada. Normalmente no es posible estudiar toda la población:
  - Destrucción de los elementos: ej. estudiar la tensión de rotura de cables.
  - Lo elementos pueden existir conceptualmente, pero no en la realidad: ej. Población de piezas defectuosas que producirá una máquina.
  - Inviable económicamente estudiar toda la población.
  - El estudio llevaría tanto tiempo que sería impracticable.
- Se suele elegir un conjunto representativo que es la **muestra**, y si esta se selecciona bien podemos obtener una información similar de la población.
- La clave es seccionar la muestra **representativa** de la población.

# Métodos de muestreo: muestreo aleatorio simple

- Una muestra es aleatoria simple (m.a.s.), si :
  - Cada elemento de la población tiene la misma probabilidad de ser elegido.
  - Las observaciones se realizan con reemplazamiento (población idéntica en todas las extracciones).
- La primera condición asegura representatividad de la muestra (si A esta en el 20% y todos los elementos tienen idéntica probabilidad de ser seleccionados, la muestra tendrá un 20% también de A).
- La segunda se impone por simplicidad.

# Métodos de muestreo: muestreo aleatorio simple

- En una muestra aleatoria cada observación tiene la distribución de probabilidad de la población (precisamente porque es aleatoria).
- Sea la muestra observada $X'=(x_1, \ldots, x_n)$, donde $x_i$ representa el valor de x en el elemento i-ésimo.
- Llamamos $f_1,\ldots,f_n$ a las funciones de densidad de esas variables que verifican en el muestreo aleatorio simple que $f_1=\ldots=f_n=f$.
- Como las observaciones son independientes en una muestra aleatoria simple, entonces la distribución conjunta de la muestra se puede poner $f_c(x_1, \ldots, x_n)= f_1(x_1)\, f_2(x_2)\ldots\, f_n(x_n) \;= f(x_1)\, f(x_2)\ldots f(x_n)$.

# Métodos de muestreo: otros tipos de muestreo

- Muestreo **estratificado**: El método anterior se utiliza cuando la población es homogénea. Cuando se tiene información heterogénea de la población hay que dividir la población en estratos o clases, realizando un muestreo aleatorio simple dentro de cada estrato (ej. Encuestas de opinión, que se divide por sexo, edad, profesión, etc.).
- Supongamos k estratos de tamaños $N_1,\ldots N_k$, con $N=N_1+\ldots N_k$.
- La muestra que tomemos debe garantizar la presencia adecuada de cada estrato.
- Existen criterios básicos para dividir o repartir el tamaño total de muestra (n) entre los estratos ($n_i$):
  - Proporcional: $n_i=n * (N_i/N)$.
  - Proporcional a la variabilidad del estrato: los estratos variables están más representados. Si $\sigma_i$ es la variabilidad del estrato i, entonces:

    $$n_i=n * (\sigma_i N_i)/(\sum^k_{i=1} \sigma_i N_i)$$

# Métodos de muestreo: otros tipos de muestreo

- Muestreo por **conglomerados**: Hay situaciones en las cuales donde ni el muestreo aleatorio simple ni el estratificado pueden darse. En estos casos la población se encuentra agrupada en conglomerados, cuyo número se conoce.
- Ej. La población se distribuye en provincias, los habitantes de provincias en ciudades, etc.
- Si se suponen los conglomerados independientes se pueden analizar con la metodología anterior.

# Métodos de muestreo: otros tipos de muestreo

- Muestreo **sistemático**: Cuando los elementos están ordenados en listas. Supongamos que queremos una muestra n de una población N. Calculamos $k=N/n$. Se coge un elemento entre los primeros k, supongamos que el orden del elegido es $n_1$, tomamos a continuación los elementos $n_1+k$, $n_1+2k$, etc., hasta completar la muestra de tamaño n, es decir n veces (siendo k el número de grupos de tamaño n en la población).
  - Si el orden de los elementos en la lista es al azar este procedimiento es equivalente al muestreo aleatorio simple.
  - Si el orden de los elementos en la lista es de la forma que elementos cercanos son más similares que los más alejados, entonces se puede demostrar que este procedimiento cubre más homogéneamente toda la población, siendo más preciso que el muestreo aleatorio simple.

# La estimación puntual: fundamentos

- **Supongamos que se observa una muestra aleatoria simple de una variable aleatoria x siguiendo una distribución conocida** como las que hemos estudiado: distribución normal, Poisson, etc.
- Lo que **no conocemos son los parámetros** de esas distribuciones conocidas.
- A las cantidades que estiman los parámetros de la distribución de la población a través de datos muestrales se le llaman **estimadores** estadísticos.
- ¿Cómo estimamos esos parámetros de los datos muestrales recogidos?

# La estimación puntual: fundamentos

- En primera aproximación supondremos que **no tenemos ningún tipo de información del parámetro a ajustar**, $\vartheta$, de la distribución supuesta.
- **Si hubiese algún tipo de evidencia sobre el parámetro** (unos valores son muchos más probables que otros) **a estimar se utiliza el enfoque bayesiano** (más adelante).
- Así el enfoque que vamos a ver ahora es el paramétrico, que dependiendo del tipo de variable a estudiar supondrá un modelo u otro a ajustar sus parámetros.
- Por lo tanto la forma del modelo se conoce en el enfoque paramétrico. El modelo se debería saber de la información previa de los datos muestrales.
- Vamos a continuación a ver algún ejemplo simple de como extraer el posible modelo de los datos muestrales.

# La estimación puntual: la identificación del modelo

- La primera operación a realizar con la muestra en un análisis descriptivo, para así ver si el modelo que consideramos es consistente con la muestra.
- Si tenemos muestras pequeñas (menos que 30), es más complicado y se suelen hacer ciertos tipos de gráficos que nos sacan de dudas.
- Por ejemplo, para chequear visualmente si una muestra pequeña la podemos asociar a una distribución de Poisson:
  - Si **siguen una distribución de Poisson** entonces $\boxed{E[f_{ob}(x)]=nP(x)=n\frac{\lambda^x}{x!}e^{-\lambda}}$, siendo n el tamaño de la muestra (i.e. la frecuencia de observar $n_r$ $=E[f_{ob}(x=r)]$ de un total de n observaciones, para la clase r será nP(x=r)).
  - Si sacamos logaritmos neperianos: ln $E[f_{ob}(x)]=\ln n - \lambda + x\ln\lambda - \ln x!$, por lo tanto ln $E[f_{ob}(x)] + \ln x! = \ln n - \lambda + x\ln\lambda$ = A+xB. (notar que ln $E[f_{ob}(x)]$=ln $f_{ob}(x)$, ya que E[c]=c, con c=cte.).
  - Por tanto si dibujamos ln $f_{ob}(x) + \ln x!$ respecto de x tiene que salir casi una recta si los valores esperados se distribuyen según una distribución de Poisson (**ver ejemplo 7.1** numérico en el libro).
  - La recta debería tener una pendiente $\ln\lambda$ y ordenada en el origen $(\ln n - \lambda)$.

# La estimación puntual: la identificación del modelo

- Existen otros métodos para comprobar otras distribuciones, como por ejemplo para la distribución normal:
  - Que se puede utilizar un papel probabilístico normal (ver ejemplo en el libro) para dibujar los datos. Sino se ajustan a una recta, los datos no se distribuyen según una normal (**ver figura 7.4 del libro**).
  - También se pueden usar los gráficos "Q-Q plots" con la misma idea anterior, que ya hemos visto anteriormente.
- En general los gráficos "Q-Q plots" se pueden utilizar con cualquier distribución de probabilidad, como ya comentamos anteriormente.

# La estimación puntual: el método de los momentos

- Es el primer método que se utilizó para obtener el ***estimador*** de un parámetro de una distribución dada (formalizado por K. Pearson).
- Recordemos que un ***estimador*** es un valor obtenido a partir de los datos muestrales para un parámetro de la distribución de probabilidad que se supone que representará a la población.
- Se toma como estimador de la varianza de la población la varianza de la muestra, de la media de la población la media muestral, y así sucesivamente con todos lo momentos que se quieran incluir en la estimación paramétrica.
- Se trata de estimar un vector de parámetros $\underline{\vartheta} = (\vartheta_1, \ldots, \vartheta_k)$, cuyos componentes se pueden expresar en función de los k momentos de la población, $m_k$, siendo $\vartheta_1 = g_1(m_1,\ldots,m_k)$, …, $\vartheta_k = g_k(m_1, \ldots, m_k)$.

# La estimación puntual: el método de los momentos

- Así estimamos los correspondientes momentos muestrales, $\widehat{\mathrm{m}}_1, \dots, \widehat{\mathrm{m}}_k$, sustituyéndolos en el sistema de ecuaciones anteriores, obteniendo los parámetros estimados en la población: $\widehat{\vartheta}_1, \dots, \widehat{\vartheta}_k$ (ver **ejemplos 7.2 y 7.3**).
- Por ejemplo supongamos la muestra (2, 4, 9, 1) de una distribución uniforme en el intervalo (0, b), para estimar b de la muestra (el valor esperado de una distribución uniforme en a, b, es a+b/2):
  - $\vartheta_1 = g_1(m_1)$, $m_1 =$ E[x] = (a+b)/2, $\vartheta_1$=b=2E[x] – a=$g_1(m_1)$,
  - $\widehat{\vartheta}_1 = \hat{b} = 2\widehat{m}_1 - a$, $\hat{b} = 2\bar{x} - a$. ($\widehat{m}_1$ es el momento muestral primero, que es la media de la muestra medida, $\bar{x}$).
  - E[x]=(0+b)/2, por tanto $\hat{b} = 2\bar{x}$=2((2+4+9+1)/4)=8, este estimador no es el más preciso, ya que pudiéramos elegir el máximo, 9. Además si tenemos otra muestra cambia.

# Distribución de un estimador en el muestreo: concepto

- Es muy importante estimar cual es la bondad de ajuste de estos estimadores de los parámetros y sus propiedades deseables.
- Cuando nos podemos fiar de estos estimadores para obtener información fidedigna de la población cuando nos basamos en parámetros de las muestras.
- Esto es lo que vamos a ver a continuación.
- **Podemos ver el estimador como una variable aleatoria**, cuyo valor cambia de muestra en muestra.
- Consideremos una población de la que se toman muestras con remplazamiento de tamaño n y calculamos en cada muestra la media $\bar{x}$.

# Distribución de un estimador en el muestreo: concepto

- Así si tomamos k muestras obtendremos en general k valores de las medias muestrales: $\bar{x}_1, \dots, \bar{x}_k$. Si k es muy grande tendiendo a infinito los valores los valores $\bar{x}_i$ tendrán una distribución que llamaremos **distribución muestral de la media** en el muestreo.
- Esta distribución en el muestreo de un estadístico depende de:
  - La población base.
  - El tamaño de la muestra n.

# Distribución de un estimador en el muestreo: concepto

- Recordemos que el **estudio** de determinadas **características** de una **población** se efectúa **a través** de diversas **muestras** que pueden extraerse de ella (no olvidar).
- En general el **muestreo** se puede realizar **con o sin reposición.**
- La población de partida puede ser infinita o finita.
- Una población finita en la que se efectúa muestreo con reposición podría considerarse infinita, aunque también una población muy grande puede considerarse como infinita.
- Aquí vamos a limitarnos a una **población** de partida **infinita** o a **muestreo con reposición.**

# Distribución de un estimador en el muestreo: concepto

- Si consideramos todas las posibles muestras de tamaño n en una población, para cada muestra se puede calcular estadísticos como la media, desviación típica, proporción, etc., que variarán de una muestra a otra.
- Así obtenemos una distribución del estadístico que consideremos que se es lo que se llama **distribución muestral**, en general.
- Un **estadístico** (refiriéndose a datos muestrales, i.e. estadístico muestral) es una medida cuantitativa, derivada de un conjunto de datos de una muestra, con el objetivo de estimar o inferir características de una población o modelo estadístico.

# Distribución de un estimador en el muestreo: concepto

- En general podemos definir un **Estadístico** como una función de los valores de la muestra. Es una variable aleatoria, cuyos valores dependen de la muestra seleccionada.
- Su distribución de probabilidad, se conoce como **Distribución muestral del estadístico**.
- Por ejemplo para el estadístico media de las diferentes muestras de una población se obtiene la **distribución muestral de la media**, para la varianza la **distribución muestral de la varianza**, etc.

# Distribución de un estimador en el muestreo: distribución en el muestreo de una proporción

- Supongamos una población donde **observamos la presencia o no de un atributo.** Y sea p la proporción desconocida de elementos con dicho atributo en la población.
- La distribución del muestreo del estimador de $\hat{p}$ en la muestra, viene determinada por la distribución binomial:
  - $P\left(\hat{p} = \frac{r}{n}\right) = PB(r) = \binom{n}{r} p^r (1-p)^{n-r}$, r=0, 1, …, n
- Por lo tanto la **probabilidad de que la proporción** en la muestra sea r/n es igual a la probabilidad de obtener r elementos con una característica determinada en una muestra de tamaño n, que es directamente la **distribución binomial.**

# Distribución de un estimador en el muestreo: distribución en el muestreo de una proporción

- Así las propiedades de la distribución en el muestreo del estimador $\hat{p}$ vendrán dadas por la propiedades del valor esperado y la varianza del estimador de la proporción:
  - **E[$\hat{p}$]**=E[r/n]=(1/n)E[r]=np/n=**p,**
    - (recordar que en una binomial E[r]=np).
  - **Var[$\hat{p}$]**=Var[r/n] =$(1/n)^2$ Var[r]= **pq/n,**
    - (recordar que en una binomial Var[r]=npq).
  - **Recordar** las propiedades de la esperanza y varianza: E[aX]=aE[X] y Var[aX]=$a^2$Var[X]
- Observar que cuando el tamaño de las muestras va a infinito, la varianza de la distribución del muestreo del estimador va a 0.

# Distribución de un estimador en el muestreo: distribución en el muestreo de una proporción

- Cuando n es grande, la distribución de muestreo de $\hat{p}$ será aproximadamente normal con la media y varianza de las dos expresiones anteriores, ya que es un caso particular de la distribución muestral de una media, ya que $\hat{p}$ se calcula por: $\hat{p}=(x_1+\ldots+x_n)/n$ y entonces se puede aplicar aplica el TCL.
  - Así $\hat{p}$ se comporta como N(p, $(pq/n)^{0.5}$) por el TCL.
- Cada $x_i$ toma el valor 1 si el elemento tiene el atributo estudiado, y 0 en cualquier otro caso.
- **Así $\widehat{p}$ es la media muestral de las variables de Bernoulli, $x_i$.**

# Distribución de un estimador en el muestreo: distribución en el muestreo de una proporción: Ejemplo

- Supongamos que una empresa de paquetería sufre una proporción del 3% de retrasos en sus envíos. Cual es la probabilidad de habiendo realizado 500 pedidos sufran mas del 4% de retrasos en los mismos:
- p=0.03, n =500, P($\hat{p}$ >0.04)????, sabemos que $\hat{p} \sim N(p, (pq/n)^{0.5})$ =N(0.03, 0.0076). Ahora estandarizamos la variable a $z=(\hat{p}-\mu)/\sigma$, para así calcularla a través de z=N(0,1).
- z=(0.04-0.03)/0.0076=1,32, así P($\hat{p}$ >0,04)=P(z>1,32)=

  1- P(z≤1,32)=1- 0,90658 = 0.0934 (miramos en la tabla 4 de las tablas al final del libro para la distribución normal estandarizada, para 1,32).

# Distribución de un estimador en el muestreo: distribución muestral de la media

- Para calcular la distribución muestral de la media tenemos que tener en cuenta que cada muestra de tamaño n que podemos extraer de una población proporciona una media.
  - Podemos considerar cada una de estas medias como valores de una variable aleatoria y podemos estudiar su distribución que llamaremos **distribución muestral de las medias.**
- Vamos a calcular la media y varianza de la distribución muestral de la media, en el caso general en el que la variable aleatoria x tiene media $\mu$ y varianza $\sigma^2$.

# Distribución de un estimador en el muestreo: distribución muestral de la media

- Para el cálculo de distribución muestral de la media suponemos que cada muestra es de tamaño n, y suponemos que todas las variables $x_i$ de una muestra aleatoria simple tiene la misma distribución de la población.
- La variable aleatoria ahora es $\bar{x}=(x_1+\ldots+x_n)/n$.
- Así el valor esperado la distribución muestral de las medias y su varianza vienen determinados por las expresiones:
  - $E[\bar{x}]=E[(1/n)\sum x_i]=1/n \sum E[x_i]= 1/n \sum \mu= (1/n)\, n\mu = \mu$.
  - $Var[\bar{x}]=(1/n)^2 \sum Var[x_i]= n\sigma^2/n^2= \sigma^2/n$.
- Hemos aplicado el hecho de que para la variable aleatoria $x_i$ se cumple que $E[x_i]=\mu$ y $Var[x_i]= \sigma^2$, como hemos dicho antes.

# Distribución de un estimador en el muestreo: distribución muestral de la media

- Resumiendo, al tomar una muestra de tamaño n de una variable con media μ y varianza $\sigma^2$ y distribución cualquiera, la distribución muestral de la media verifica que E[$\bar{x}$]=μ y Var[$\bar{x}$]= $\sigma^2$ /n.
- En el caso de la distribución de muestreo de la proporción E[$\hat{p}$]=p y Var[$\hat{p}$]=pq/n, es un caso especial de este que acabamos de ver con media p y varianza pq.
- Si tenemos una población normal N(μ, σ) y extraemos de ella muestras de tamaño n, la distribución muestral de medias sigue también una distribución normal N(μ, $\sigma/(n)^{0.5}$) (volver). Darse cuenta que $\bar{x}$ es una combinación lineal de variables aleatorias normales.
- Notar que cualquier combinación lineal de variables aleatorias normales independientes es una variable aleatoria normal con media la misma combinación lineal de las medias y con varianza la combinación lineal de las varianzas con los coeficientes que las acompañan al cuadrado (ver sección 6.8, apartado 5) del libro).
- **Si la población no sigue una distribución normal** pero n>30, aplicando el llamado el TCL la distribución muestral de medias se aproxima también a la normal anterior (razonamiento similar al de la proporción, 3 transparencias atrás, ya que $\bar{x}$ se calcula por: $\bar{x}=(x_1+\ldots+x_n)/n$ y entonces se puede aplicar el TCL).

# Distribución de un estimador en el muestreo: distribución muestral de la media: Ejemplo

- Supongamos que el diámetro en mm de la fabricación de unos dulces se distribuye según una normal N(167, 3,2). Cual es la probabilidad de comprado 10 de esos dulces, la media de los diámetros de los mismos sea menor o igual que 165 mm:
- $\mu$=167, $\sigma$=3,2, n =10, P($\bar{x}$ $\leq$165) ????, sabemos que $\bar{x}$~N($\mu$, $\sigma/(n)^{0.5}$) = N(167, 1,012). Ahora estandarizamos la variable a z=($\bar{x}$-$\mu$)/$\sigma$, para así calcularla a través de z=N(0,1).
- z=(165-167)/1,012=-1,98, así P($\bar{x}$ $\leq$165)= P(z$\leq$-1,98)
- La tabla 4 de las tablas al final del libro para la distribución normal estandarizada solo contiene los datos para valores positivos de *z*, ya que, por simetría:
  - $F(-|z_0|) = P(z \leq -|z_0|) = P(z \geq |z_0|) = 1 - F(z_0)$.
- Así P($\bar{x}$ $\leq$-1,98)=1- P($\bar{x}$ $\leq$1,98)=1- 0,97615 =0,02385 (miramos en la tabla 4 de las tablas al final del libro para la distribución normal estandarizada).

# Distribución de un estimador en el muestreo: distribución muestral de la varianza

- Para la **distribución muestral de varianzas**, podemos seguir los mismos razonamientos anteriores y suponemos de nuevo una variable aleatoria x que se observa en la muestra que tiene media μ y varianza $\sigma^2$.
- Se puede calcular que esperanza de la distribución de varianza de la muestra es $E[s^2]=[\sum(x_i-\mu)^2/n]=\sigma^2(n-1)/n$. En consecuencia el valor esperado de $s^2$ es menor que $\sigma^2$, aunque la diferencia tiende a cero a aumentar el tamaño de la muestra n.
- Se puede definir la varianza muestral corregida como $\hat{s}^2=[\sum(x_i-\mu)^2/(n-1)]=(n/n-1)\ s^2$, de tal forma que $E[\hat{s}^2]=\sigma^2$.
- Estas propiedades se verifican siempre, cualquiera que sea la distribución de la variable x.
- Así se pueden calcular la distribución de cualquier estimador en el muestreo, en libro vienen más ejemplos (capítulo 7).

# Propiedades de los estimadores: centrado o insesgado

- Diremos que un **estimador** $\hat{\vartheta}$ es **centrado** o in-sesgado para $\vartheta$ si para cualquier tamaño muestral tenemos que E[$\hat{\vartheta}$]= $\vartheta$.
- Cuando no es centrado se define el **sesgo** del estimador como: $sesgo\left(\hat{\vartheta}\right) = E\left[\hat{\vartheta}\right] - \vartheta.$
- **Pueden existir muchos estimadores centrados para un parámetro:**
  - Por ejemplo, para estimar μ en una distribución cualquiera todos los estimadores del tipo siguiente son centrados: $\hat{\mu} = a_1x_1 + \cdots + anxn$ con $\sum a_i = 1.$
- Anteriormente hemos comprobado que $\bar{x}$ (y como caso particular el estimador de la proporción $\hat{p}$) es siempre centrado para estimar μ.

# Propiedades de los estimadores: centrado o insesgado

- También hemos visto que $s^2$ no es centrado para estimar $\sigma^2$ (recordar que lo corregimos).
- Una ventaja fundamental de los **estimadores centrados** es que **los podemos combinar para obtener nuevos estimadores centrados:**
  - Si tenemos dos muestras independientes y calculamos en cada una de ellas un estimador centrado $\hat{\vartheta}_i$ para un para un parámetro determinado, cualquier estimador del tipo $\hat{\vartheta}_t = a_1\hat{\vartheta}_1 + a_2\hat{\vartheta}_2$, con $a_1 + a_2 = 1$, es un estimador centrado.
- Los **estimadores centrados no tienen porque ser los mejores**, alguna veces es preferible tener uno sesgado con poca varianza, que uno no sesgado pero con mucha varianza (**eficiencia de estimadores**).

# Propiedades de los estimadores: eficiencia o precisión

- La **eficiencia** o precisión se define en función del **inverso de la varianza**: $\text{precisión}\left(\hat{\vartheta}\right) = 1/Var\left[\hat{\vartheta}\right]$.
- Diremos que un estimador $\hat{\vartheta}_2$ es **más eficiente** que un estimador $\hat{\vartheta}_1$ si para cualquier tamaño muestral se cumple que $\text{Var}\left(\hat{\vartheta}_2\right) \leq \text{Var}\left(\hat{\vartheta}_1\right) \Leftrightarrow \text{Efic}\left(\hat{\vartheta}_2\right) \geq \text{Efic}\left(\hat{\vartheta}_1\right)$.
- Se define la **eficiencia relativa** de $\hat{\vartheta}_2$ respecto a $\hat{\vartheta}_1$ al cociente entre sus eficiencias: $\text{ER}\left(\hat{\vartheta}_2/\hat{\vartheta}_1\right) = \text{Efic}\left(\hat{\vartheta}_2\right)/\text{Efic}\left(\hat{\vartheta}_1\right) = \text{Var}\left(\hat{\vartheta}_1\right)/\text{Var}\left(\hat{\vartheta}_2\right)$ .
- La **eficiencia** de estimadores esta completamente ligada a la **varianza** de los mismos.
- Mirar en el libro como combinar linealmente estimadores centrados (sección *Combinación lineal de estimadores centrados* de la sección 7.5.2) para minimizar la varianza y el ejemplo 7.4.

# Propiedades de los estimadores: error cuadrático medio

- Muchas veces se nos presenta el problema de elegir entre dos estimadores uno centrado y con varianza no muy grande y otro sesgado y con varianza un poco más pequeña. En estos casos se elige aquel que tiene el menor error cuadrático medio con el parámetro que esta intentando estimar.
- Así tenemos $\mathrm{ECM}(\vartheta) = E\left[\left(\hat{\vartheta} - \vartheta\right)^2\right]$, tomando el promedio respecto a la distribución en el muestreo del estimador.
- Se puede demostrar que $\mathrm{ECM}(\vartheta) = \left[sesgo\left(\hat{\vartheta}\right)\right]^2 + Var\left(\hat{\vartheta}\right)$.

# Propiedades de los estimadores: consistencia y robustez

- Cuando lo único que podemos tener son estimadores segados y no con mucha eficiencia, se le pide al estimador que sea **consistente.**
- Un estimador se dice que es consistente si cuando crece el tamaño de la muestra el estimador tiende al parámetro que se está estimando: $\lim_{n\to\infty} E[\hat{\vartheta}_n] \to \vartheta$.
- Es decir la esperanza del estimador es asintóticamente el valor del parámetro.

# Propiedades de los estimadores: consistencia y robustez

- En este caso la varianza del estimador va a cero también con el tamaño de la muestra: $\lim_{n\to\infty} Var[\hat{\vartheta}_n] \to 0.$
- Un buen estimador es **robusto** para un parámetro $\vartheta$ en el modelo f(x), si variando débilmente el modelo este estimador experimenta una pequeña modificación (ver ejemplo en el libro, capítulo 7 de contaminación de un modelo normal como decrece la eficiencia).