



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

ASTROMETRÍA DESDE UN ENFOQUE BAYESIANO

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL ELÉCTRICO

ALEX MAURICIO ECHEVERRÍA SOLÍS

PROFESOR GUÍA:
JORGE SILVA SANCHEZ

MIEMBROS DE LA COMISIÓN:
RENÉ MENDEZ BUSSARD
MARCOS ORCHARD CONCHA

SANTIAGO DE CHILE
2016

**RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE:** Ingeniero Civil Eléctrico
POR: Alex Mauricio Echeverría Solís
FECHA: Santiago, 2016
PROFESOR GUÍA: Jorge Silva Sanchez

Astrometría desde un enfoque bayesiano

En la Astronomía ha habido un salto cuantitativo gigantesco desde el nacimiento de la tecnología CCD y las imágenes digitales. A pesar de ello, todavía existe un espacio de mejora en lo que respecta a las técnicas para estimación de parámetros importantes que caracterizan a las estrellas. Es por eso que esta Memoria de Título se presenta como objetivo el estudiar y cuantificar el uso de nuevos enfoques de estimación modernas no aplicados aún en esta disciplina para la estimación de la posición de objetos luminosos (Astrometría).

Para poder entender el problema se presenta qué es una cámara digital y su uso en la astronomía, específicamente en la astrometría, además de presentar importantes conceptos astronómicos que se usan a lo largo de la memoria, como lo son el *Point Spread Function* y el *Full Width at Half Maximum*. Por otro lado, se da un repaso a los elementos de estimación necesarios para resolver el problema, como Cramér-Rao, Cramér-Rao Bayesiano y los estimadores Esperanza Condicional, *Maximum Likelihood* y *Least Squares*.

La implementación del estimador se realizará a partir de una formalización completa del problema de estimación en astrometría, donde se incluirá también el trabajo de los algoritmos necesarios para encontrar el valor numérico tanto del estimador como de su error cuadrático medio. Se mostrará también la resolución de la cota de Cramér-Rao, tanto para la versión paramétrica como la bayesiana.

Se hace un análisis de las herramientas presentadas usando como figura de mérito el MSE (Error Cuadrático Medio). A partir de ello, se muestra cómo varía este valor como función del tamaño del pixel, la relación de señal-ruido y sus ganancias relativas, para posteriormente estudiar las diferencias entre la Cota Bayesiana de Cramér-Rao y el MSE de la Esperanza Condicional, el estimador propuesto para el problema.

Finalmente se concluye, viendo que existen ganancias significativas del enfoque Bayesiano en regímenes de baja relación señal-ruido y gran tamaño de pixel. Además se verifica que la cota Bayesiana de Cramér-Rao es un buen predictor del MSE de la Esperanza Condicional, lo cual trae nuevas preguntas que se pueden plantear como trabajo futuro a partir de esta Memoria de Título.

*A mi familia y mis compañeros de vida, que a través del andar
forjaron lo que soy y seré*

Agradecimientos

Estas páginas recogen alrededor de dos años de experiencias y aprendizajes que muchas veces fueron más allá de lo técnico o lo científico. Este tiempo de constante crecimiento ha sido posible solamente por todas esas personas que me han aportado en pensar y actuar de una forma distinta, es por eso mismo que estas palabras me parecerían estériles si no las nombro.

Agradecer primero a Jorge Silva, René Mendez y Marcos Orchard, los profesores a cargo de mi tema de investigación, que con su dedicación y preocupación me instaron a seguir adelante. También agradecer a todo el equipo del IDS: A Felipe, Thomas, Andrea, Miguel, Hernán, Roberto, Sebastian y Rodrigo, que significaron para mí más que un espacio de trabajo sino un lugar de amistad y apoyo.

A Andrés, Javier y Gonzalo, por ser con quienes empecé y terminé este viaje por la carrera de Ingeniería Eléctrica. A Carolina, Daniel, Francisca y Rodrigo, quienes conocí al andar y forjé una profunda amistad a través muchos viajes, fiestas y trabajo organizado. A Carlo y Nicolas, con quienes hemos buscado en nuestro ideales la fuerza para construir nuevas formas de organización social. A mis amigos mechones y de colegio: Eial, Hernán, Eduardo, Camila, Patricio, que si bien los últimos años nos fuimos distanciando, siempre supe que podía contar con ellos. Finalmente a Claudia, mi compañera de vida, que ha sido un soporte fundamental el último tiempo. Agradecerles a todos su incondicional amistad y apoyo en todo momento.

Este trabajo se lo dedico a mi familia, que ha ido creciendo desde que entré a la Universidad. Agradecerle al destino por conocer a mi prima/sobrina Arielle, con quién he compartido tan grandes momentos y aprendizajes, además de permitirme conocer a su familia, y de paso, hacerla parte de la mía. A mis primas Claudia y Constanza que a pesar de la distancia, siempre estuvieron ahí. Finalmente agradecer a mi querido padre por su fuerza y su determinación abnegada a cuidarnos. A mi hermana Nathalie, mi pequeña sabia de dulzura incalculable, y mi hermano Alvaro, que con su madurez y compañía me han salvado de los momentos más difíciles.

Cuando era pequeño mi Madre siempre me pedía dos cosas: Que fuera feliz y que confiara en mi mismo. En este nuevo ciclo como profesional, tomaré esas palabras...

Alex Echeverría Solís

Tabla de contenido

Agradecimientos	III
1. Introducción	1
1.1. Motivación	1
1.2. Definición del Problema	2
1.3. Objetivo General	2
1.4. Objetivos Específicos	3
1.5. Hipótesis	3
1.6. Alcances	4
1.7. Estructura de la memoria	4
2. Marco Teórico	5
2.1. La Cámara CCD y adquisición de imágenes	5
2.2. Astrometría	7
2.2.1. Términos astronómicos relevantes	8
2.3. Elementos de Estimación	11
2.3.1. El principio Bayesiano	11
2.3.2. La esperanza condicional	13
2.3.3. Estimación Paramétrica	13
2.3.4. La Cota de Cramér-Rao	15
2.3.5. La Cota de Cramér-Rao Bayesiana (Desigualdad de Van Trees) [10]	15
2.3.6. Comparación entre estimación paramétrica y Bayesiana	16
3. Implementación	18
3.1. Descripción del problema	18
3.1.1. Astrometría en CR y BCR	21
3.2. Estimador Bayesiano del Problema	23
3.2.1. Caso Astrométrico	23
3.3. Implementación de la Esperanza Condicional en Astrometría	26

3.4. Implementación de la Varianza Condicional en Astrometría	29
4. Análisis	32
4.1. Análisis Numérico	32
4.1.1. MSE con respecto al tamaño del pixel	32
4.1.2. Ganancia Relativa del caso bayesiano y aporte de la distribución a priori . . .	34
4.1.3. Brillo equivalente de los objetos	36
4.2. Resultados experimentales	38
4.2.1. Caso p.d.f a priori Uniforme	38
4.2.2. Caso p.d.f Normal	38
5. Conclusiones y trabajos futuros	42
Bibliografía	44

Índice de figuras

2.1. CCD	6
2.2. CCD y adquisición serial de datos	7
2.3. Estrella en un sensor CCD	8
2.4. Ejemplo de patrón de una estrella en el límite de difracción	9
2.5. Ejemplo del patrón de una estrella con distribución normal	9
2.6. Full Width at Half Maximum (FWHM)	10
3.1. Medición del arreglo unidimensional de píxeles	19
3.2. Discretización de una estrella en un arreglo unidimensionales de píxeles	19
3.3. Funcionamiento simplificado de un arreglo de píxeles como un integrador de fotones	21
3.4. Matriz de datos de λ_{i,x_c} , donde cada fila completa contiene a una estrella de acuerdo al modelo planteado, mientras que por cada columna se encuentra la posición de esta, que es el parámetro a estimar	26
3.5. Creación de una realización en virtud de la variable aleatoria de x_c	30
3.6. MSE acumulado en función de la cantidad de iteraciones, para un $FWHM = 1''$, $SNR = 6$, $fs = 2000[ADU/arcsec]$, $D = 0$, $RON = 5 [e^-]$	31
4.1. Error cuadrático medio en función del tamaño del pixel, para distintos $FWHM$. En ambos casos la información a priori es normal de $\sigma_{priori} = 0.5''$. La línea punteada corresponde a valores de BCR y la línea entrecortada a valores de MCR.	33
4.2. Error cuadrático medio en función del tamaño del pixel, para distintos $FWHM$. En ambos casos la información a priori es normal de $\sigma_{priori} = 0.1''$. La línea punteada corresponde a valores de BCR y la línea entrecortada a valores de MCR.	33
4.3. Error cuadrático medio en función del tamaño del pixel, para distintos $FWHM$. En ambos casos la información a priori es normal de $\sigma_{priori} = 0.05''$. La línea punteada corresponde a valores de BCR y la línea entrecortada a valores de MCR.	34

4.4. Ganancia relativa entre BCR y MCR, en función de σ_{priori} y el otro en la ganancia relativa de BCR con respecto al σ_{priori} en función del σ_{priori} , para $\Delta x = 0.2 [arcsec]$ y $FWHM = 1''$	35
4.5. Ganancia relativa entre BCR y MCR, en función de σ_{priori} y el otro en la ganancia relativa de BCR con respecto al σ_{priori} en función del σ_{priori} , para $\Delta x = 0.7 [arcsec]$ y $FWHM = 1''$	35
4.6. Brillo aparente de objetos, en relación a $\frac{F}{F_{par}}$ con respecto a σ_{priori}	37
4.7. MSE v/s tamaño del pixel, para MSE de la EC y BCR con $FWHM = 1''$ y distintos valores de SNR y σ_{priori}	39

Índice de Tablas

4.1. Valores de MSE de distintos estimadores y cotas, para distintos valores de flujo. En este caso, tanto la Esperanza Condicional como Cramér-Rao Bayesiano usan una p.d.f a priori uniforme.	38
4.2. MSE en Análisis de Flujo aparente F_{par} , para la Cota de Cramér-Rao Bayesiana y el MSE de la Esperanza Condicional, con un $FWHM = 1''$	40
4.3. MSE en Análisis de Flujo aparente F_{par} , para la Cota de Cramér-Rao Bayesiana y el MSE de la Esperanza Condicional, con un $FWHM = 2''$	40

Capítulo 1

Introducción

1.1. Motivación

La astronomía es una ciencia que tiene mucho años y es quizá una de las primeras que empezó a estudiar el hombre, que desde tiempos inmemoriales ha estado viendo el cielo y cuestionándose la razón de la existencia de los objetos luminosos que se ven todas las noches en el firmamento. Esta ciencia se ha basado fuertemente en la medición de estos objetos en el cielo, de manera de inferir a través de La Dinámica el comportamientos de estos.

Con la llegada de la electrónica y creación de los sensores fotosensibles que posteriormente dan pie a la construcción de cámaras digitales, se abre toda una nueva forma de adquirir información del cielo, y especialmente cómo poder medir la posición de los objetos en el espacio.

Si bien existe un sostenido avance de la tecnología para poder obtener imágenes astronómicas, se puede apreciar que esto no ha estado a la par con las herramientas matemáticas que se han desarrollado en el último tiempo para poder estimar parámetros de un sistema, las cuales podrían mejorar significativamente la forma en que se encuentran las posiciones de los objetos brillantes en las imágenes astronómicas. En ese sentido, desde hace varias décadas se ha estado usado un enfoque que buscan aprovechar toda la información previa que se tenga de las características del sistema para mejorar la detección o la estimación de parámetros, dicho enfoque se basan en el principio Bayesiano.

La incorporación de información previa o *a priori* de las características es algo que no se usa mucho en la Astronomía actualmente, pero que puede tener grandes repercusiones en la precisión con las que se podrían estimar objetos, ya que existe una cantidad no menor de bases de datos y catálogos ya disponibles para los astrónomos, los cuales podrían usarse para refinar sus cálculos de manera eficiente.

El tratamiento Bayesiano de este problema clásico de determinación de la posición de estrellas y objetos luminosos tiene beneficios que van más allá del mero uso de la información a priori, pues también existen garantías teóricas para alcanzar la máxima precisión en términos del Error Cuadrático Medio (que es una medida de error de estimación) al usar la Esperanza Condicional como estimador.

La propuesta es entonces utilizar las herramientas matemáticas que actualmente están en la frontera del conocimiento para aplicarlas en la estimación de la posición de estrellas en una imagen astronómica, mejorando los resultados que hoy en día se obtienen al realizar la misma tarea con herramientas clásicas de estimación.

1.2. Definición del Problema

Se tiene una imagen astronómica, la cual está compuesta por muchos objetos de luminosidad y formas de distribución de la luz variables, a la cuales es necesario determinar su posición central o *centroide*. Actualmente esto se obtiene a través de los estimadores *Least-Squares* o *Maximum Likelihood*, los cuales son estimadores paramétricos que no incorporan información previa de las características del sistema para obtener resultados. En esta Memoria de Título por tanto se realizará un análisis cuantitativo de las ganancias que se tendrían en incorporar información a priori en un sistema astronómico. Este análisis se realizará en término de los límites fundamentales para ambos enfoques de estimación; la cota de Cramér-Rao y la Cota de Cramér-Rao Bayesiana [10].

1.3. Objetivo General

Cuantificar las ganancias de utilizar el enfoque bayesiano de estimación con respecto al enfoque paramétrico en la determinación de la posición del centroide de estrellas, para el caso de un arreglo unidimensional de píxeles.

1.4. Objetivos Específicos

1. Estudiar las ecuaciones que modelan el comportamiento de las estrellas y su adquisición por parte de un arreglo de pixeles.
2. Entender los trabajos previos realizados en la determinación de cotas mínimas de Varianza en estimadores paramétricos.
3. Implementar algoritmos para determinar de las cotas de Cramér-Rao, Cramér-Rao Bayesiano y el estimador óptimo desde el punto de vista del Error Cuadrático Medio: La Esperanza Condicional.
4. Analizar las cotas de Cramér-Rao y Cramér-Rao Bayesiano, en función de los parámetros que las definen
5. Comparar los errores cuadráticos de cada enfoque en distintos regímenes y contrastarlas con el estimador Esperanza Condicional.

1.5. Hipótesis

Un argumento fuerte para empezar a analizar estos estimadores Bayesianos es que dan la posibilidad de introducir información de la características del parámetro, además de dar garantías teóricas de alcanzar la mínimo Error Cuadrático Medio (MSE), como lo que ocurre con el estimador *Esperanza Condicional*.

Sumado a esto, los trabajos de Van Trees, de Grill [3] y de Bobrovsky [1] dan muestras de la existencia de cotas para los errores cuadráticos medios de los estimadores Bayesianos, llamándolo *Cota Bayesiana de Cramér-Rao* (C-R) o simplemente desigualdad de Van Trees (V-T). Esto que abre la posibilidad de tener, además de un estimador que tiene el mínimo MSE, un predictor de su error, tal cual lo ha realizado el trabajo de Mendez et al [6].

A través del siguiente trabajo se pretender explorar estas ideas, desarrollando expresiones teóricas que se compararán en distintos regímenes, lo que llevará posteriormente a un análisis con simulaciones donde se pretende encontrar patrones de comportamiento similares entre las cotas teóricas los valores del MSE del estimador Esperanza Condicional.

1.6. Alcances

En esta memoria se estudiará el impacto del uso de la información a priori en la determinación de la posición de estrellas en arreglo de pixeles. Esto implica que la estrella se reducirá a un objeto unidimensional, el cual está caracterizado por los parámetros que definen a una Gaussiana, con una desviación estandar definida y una media que coincide con el centroide del objeto medido.

Las dinámicas propias de la adquisición de fotones por parte del sensor serán simplificadas para modelar solo los aspectos relevantes de la discretización y adquisición de la señal, pero este contendrá todos los elementos para hacer extensible este modelamiento tanto en complejidad como en dimensión.

El análisis se centrará en los resultados de la cuantificación de las diferencias de las cotas fundamentales del caso paramétricos (la cota de Cramér-Rao) y el caso Bayesiano (la cota Bayesiana de Cramér-Rao), en relación al tamaño del pixel, la relación señal-ruido (SNR) y el ancho máximo a media altura (FWHM), usando como medida de desempeño el Error Cuadrático Medio (MSE). Con respecto al análisis de estimador, este se reducirá solamente al estimador Bayesiano óptimo (la Esperanza Condicional), la cual será medida bajo distintos valores de señal-ruido.

Tanto cotas como estimadores serán medidos en base a simulaciones, controlando completamente todos los parámetros estudiados.

1.7. Estructura de la memoria

Ya habiendo introducido el tema en el Capítulo 1 de la memoria, en el Capítulo 2 se procederá a hacer una pequeña revisión bibliográfica de lo que es la cámara CCD, sus aplicaciones en Astrometría y finalmente un análisis con los estimadores que son usados en el área de la astronomía como también el estimador propuesto para la resolución del problema.

Para el Capítulo 3, se realizará un análisis de la implementación del estimador Esperanza Condicional, tanto del estimador mismo como también de la varianza condicional, desde la formalización matemática del problema hasta su implementación computacional.

Finalmente en el Capítulo 4, usando las herramientas propuestas y sus cotas fundamentales, se realizará un análisis de cómo se comporta el Error Cuadrático Medio en función del tamaño del pixel y el *Signal-to-Noise ratio* (SNR). En un último análisis se revisará cuán cercano está la Esperanza Condicional a su cota Bayesiana.

El Capítulo 5 cierra el trabajo, exponiendo las conclusiones y el trabajo futuro a realizar.

Capítulo 2

Marco Teórico

En este capítulo se realizará una revisión bibliográfica enfocada a las herramientas de estimación, tanto las usadas en el área como las propuestas en la memoria, además de una breve conceptualización de donde se usan estas herramientas y su importancia en el área.

2.1. La Cámara CCD y adquisición de imágenes

La cámara CCD es un dispositivo esencial de la astronomía. A través de este sensor se ha podido obtener imágenes del cielo nocturno, tanto de lo que ocurre en el sistema solar como también el cielo profundo, estudiando a través de la detección de luz de los más reconditos lugares, incluso donde comenzó el universo.

Tal es la importancia de las cámaras CCD, que casi no existe la astronomía moderna sin ellas. Si bien la digitalización de la información ha tenido un avance sostenido en el tiempo, no así en las herramientas matemáticas que se usan actualmente para inferir información sobre ellos. Tal como se ha dicho en el capítulo anterior, esa es la principal motivación de este trabajo.

Para poder modelar el problema de estimación de estrellas u objetos brillantes, es necesario entender cómo funciona una cámara en su función de sensor la luz.

La cámara CCD es un circuito integrado que es sensible a un rango de frecuencias de luz, el cual (en su forma más simple) se divide en una red de cuadrículas denominadas *pixeles*. Esta unidad fundamental es la encargada de generar una diferencia de potencial de acuerdo a la cantidad fotones incidentes en él, la que posteriormente se puede digitalizar y guardar en una memoria o registro computacional.

Las cámaras CCD tienen una variedad de tamaños y aplicaciones en las que se pueden utilizar, que pueden ser desde la fotografía hasta la aplicaciones científicas y de investigación.

Esquemáticamente, tendremos el funcionamiento de un sensor CCD es de la siguiente forma:

1. El sensor “recoge” fotones y los mantiene en cada pixel como una diferencia de potencial hasta

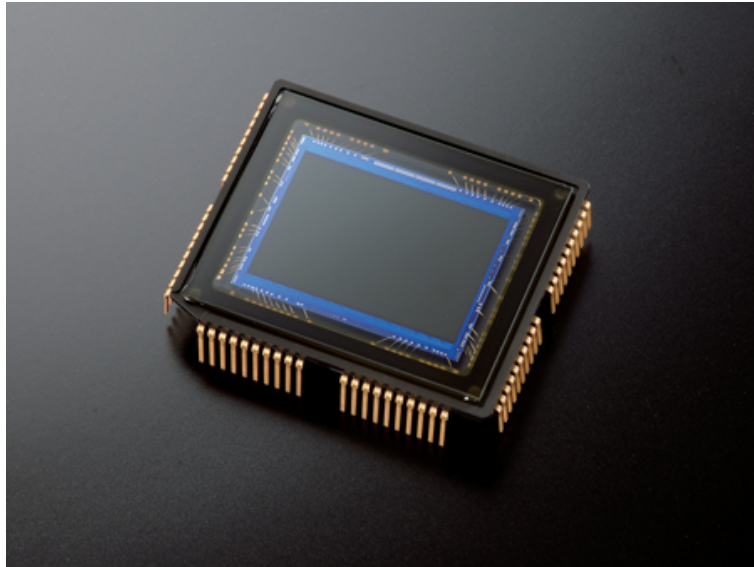


Figura 2.1: CCD¹

el momento que el sensor deja de ser expuesto a un elemento de radiación lumínica.

2. Después de ello, se inicia la lectura serial de los pixeles. Esto se realiza a través del traspaso de cargas de pixel a pixel a través de las filas, para posteriormente realizar la lectura de los pixeles a través de las columnas
3. Al leer cada pixel de cada columna se mide su voltaje, lo que finalmente se transforma en un valor discreto que posteriormente es guardado en la memoria de un computador. El proceso de conversión análogo digital genera ruido, el cual aumenta mientras más rápido se haga el proceso. El valor de salida de esta conversión es el ADU [Analog-to-Digital Unit].

La conversión análogo-digital está limitada por la cantidad de bits que pueden guardar el valor. Por ejemplo, un sistema de adquisición de 8 bits puede representar un número de tamaño $2^8 = 256$, mientras que un sistema de 16 bits puede aumentar el rango de valores a $2^{16} = 65.536$. Así, con un sistema de 16 bits puede codificarse a lo más $65.536[ADU]$ posibles niveles en un pixel. Para aplicaciones astronómicas esto es un problema, ya que la profundidad de bits no podría medir estrellas en las cuales se encuentren valores mayores, lo que es muy común en este tipo de aplicaciones. Es por esto que se establece en el arreglo de pixeles un valor de ganancia G , que aumenta el rango dinámico del sensor hasta los valores que es necesario medir, pero en desmedro de la calidad de la medición, aumentando la relación señal-ruido.

¹Marduk Astronomía. (2011), Cámaras CCD. Recuperado de <http://www.pawean.com/MVM/images/ccdchip.jpg>

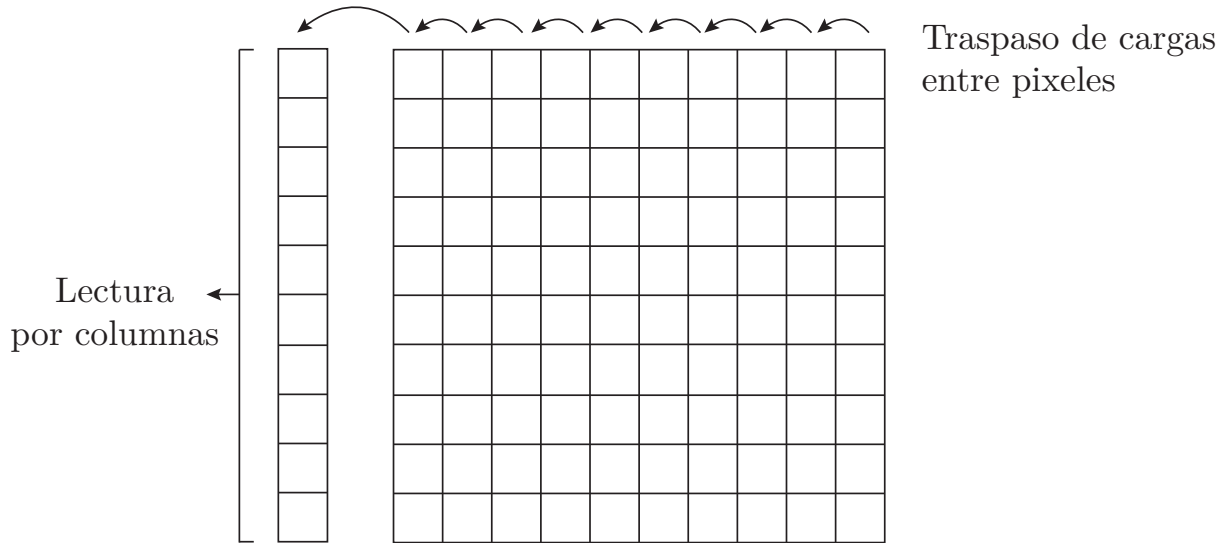


Figura 2.2: CCD y adquisición serial de datos

2.2. Astrometría

La astronomía, hasta el siglo XIX, se basó únicamente en la medición de la posición de los astros en el cielo. Estas se midieron de acuerdo a cada época histórica, con cierta exactitud limitandose a la tecnología que existía en cada momento. El estudio de las posiciones de las estrellas se le ha conocido con el nombre de *Astrometría*, y junto con el estudio del flujo (o cantidad de luz) son las formas más comunes en que los astrónomos pueden caracterizar los sistemas que están observando. Al observar la posición de los objetos, se pueden analizar el comportamiento cinemático de los objetos brillantes, pudiendo determinar por ejemplo el origen de formaciones estelares hasta sistemas más complejos como galaxias o conglomerados de estos. Esta técnica es particularmente importante en el caso del descubrimiento de exoplanetas, donde se pueden determinar la existencia de objetos opacos estudiando los cambios de posición de la estrella en la que están orbitando estos objetos. La astrometría es especialmente útil para los astrofísicos, donde puede estudiar fenómenos complejos como las dinámicas producidas por materia oscura.

En términos concretos, al observar una estrella no se ve un objeto puntual sino más bien un elemento distribuido en la imagen, el cual puede ser modelado de cierta forma (como veremos más adelante). Este objeto, de acuerdo a la modelación realizada, se le calculará a través de astrometría su *centroide*. Es este punto donde se realizarán todas las mediciones sobre el objeto brillante.

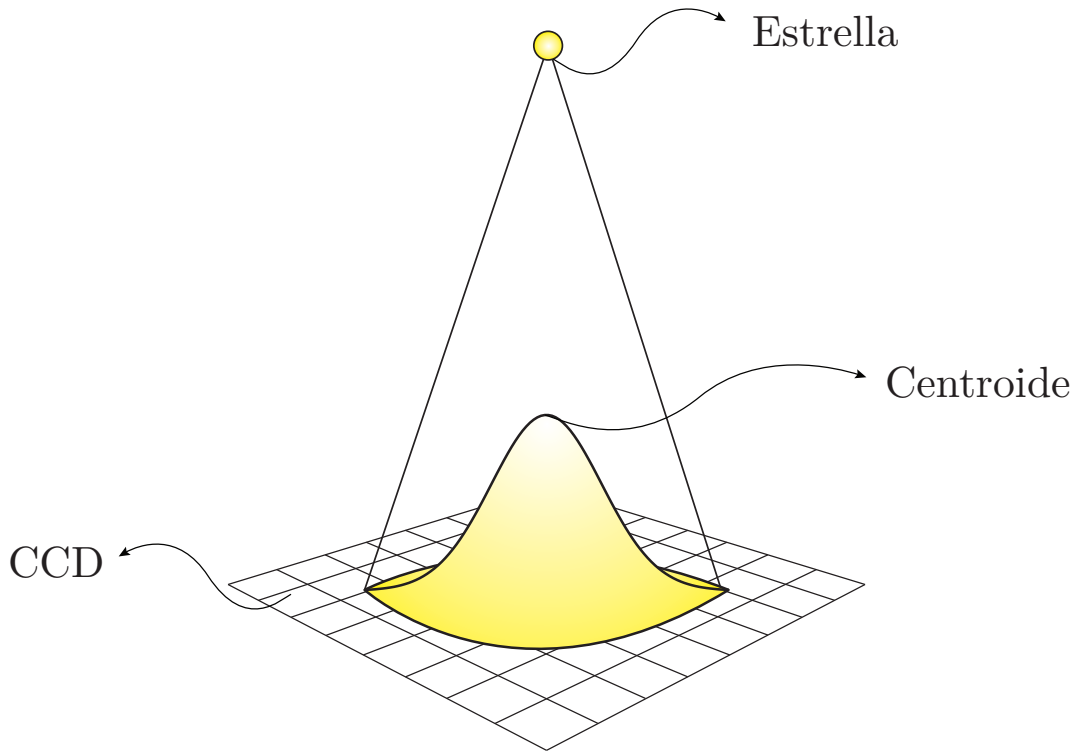


Figura 2.3: Estrella en un sensor CCD

2.2.1. Términos astronómicos relevantes

Point Spread Function (PSF)

La función de dispersión (o PSF por sus siglas en inglés) es la respuesta impulsiva de un sistema óptico enfocado. Dicho en términos astronómicos, la PSF describe como la luz se distribuye en el plano focal del telescopio, el cual posteriormente será medido en un arreglo bidimensional de píxeles.

Saber la forma de la PSF y como se comporta es fundamental, tanto en telescopios espaciales como también ubicados en tierra, donde se utilizan sistemas ópticos-adaptativos que tratán de reducir los efectos atmosféricos en la forma de la estrella.

En los casos en que no existe ningún tipo de contaminación en la señal de la estrella, la distribución observada de la estrella está solamente afectada por la difracción de la luz en el espacio. En tal caso, la estrella tiene una forma dada por el función de Bessel de primer orden, el cual contiene regiones de interferencia constructiva y destructiva (regiones iluminadas y oscuras),

Debido a diversos efectos, tanto terrestres como extraterrestes, el común de los casos es que no se pueda tener un patrón de estrella en el límite de difracción sino más bien estrellas de diversas formas y tamaños, tal como se puede ver abajo.

Para diversas aplicaciones, una buena primera aproximación es asumir que el flujo de la estrella

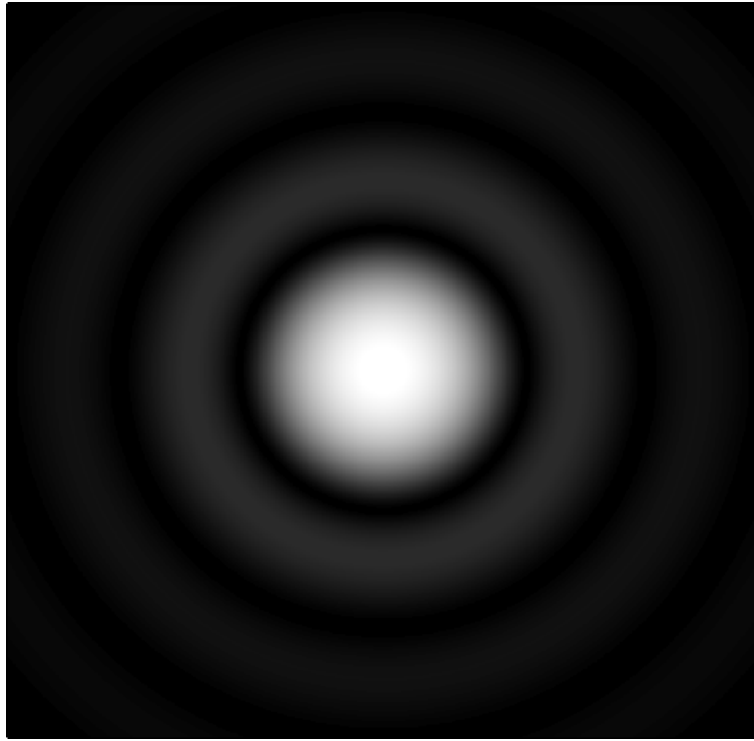


Figura 2.4: Ejemplo de patrón de una estrella en el límite de difracción

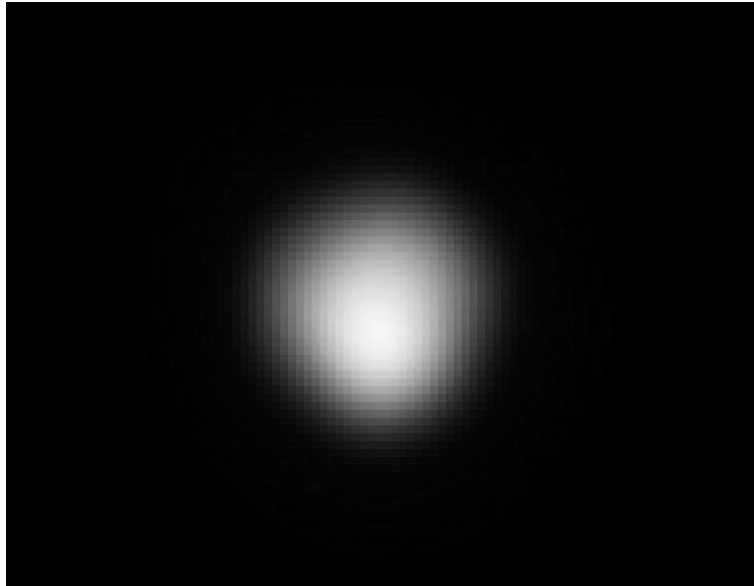


Figura 2.5: Ejemplo del patrón de una estrella con distribución normal

distribuye como una Gaussiana normalizada. En tal caso, la estrella de forma

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-x_o)^2}{2\sigma^2}\right], \quad (2.1)$$

tendrá como x_o al centroide de la estrella.

2.3. Elementos de Estimación

La teoría de estimación es una rama de la Estadística que tiene por objetivo inferir los valores de parámetros a través de mediciones como también de variables aleatorias.

Un modelo general del problema de estimación tiene las siguientes características:

1. Espacio de parámetros: Conjunto de valores posibles a estimar.
2. Función de probabilidad.
3. Regla de Estimación: Es el estimador $\hat{\theta}$ que se genera en función del enfoque dado al problema.

En un primer enfoque, el parámetro es una variable aleatoria donde su comportamiento está gobernado por una función de densidad de probabilidad. En el segundo enfoque, el parámetro es una cantidad desconocida pero no es aleatoria. Ambas escuelas tienen sus propios estimadores, agrupados en los que usan el *Principio Bayesiano* o el *Principio de Máxima Verosimilitud (Maximum Likelihood)*, respectivamente.

En esta oportunidad, se tratará de construir ideas globales en función del principio Bayesiano, para posteriormente contrastarlas con la estimación paramétrica, y en función de ello, mostrar los estimadores y las propiedades fundamentales que se tomarán en consideración en este reporte.

2.3.1. El principio Bayesiano

La idea fundamental de este enfoque es estimar el parámetro como función de la información de la medición y de la densidad de probabilidad del parámetros (ya que en este caso son una variable aleatoria).

Función de costo y riesgo

Existen ciertos funcionales necesarios para aplicar la estimación Bayesiana, uno de ellos es la función de costo.

Sea $\Theta, \Delta, D \subseteq \mathbb{R}$, la función de costo $L : \Theta \times \Delta \rightarrow \mathbb{R}$ es la pérdida o el “costo” de haber estimado δ cuando el verdadero valor es θ , por lo que $\delta = \phi(x)$ es el estimador de θ (con $\phi : X \rightarrow \Delta$).

La pérdida promedio con respecto a los datos X viene dado por la función de riesgo $\hat{R} : \Theta \times D \rightarrow \mathbb{R}$

$$\hat{R}(\theta, \phi) = \mathbb{E}[L(\theta, \phi(x))] = \int_{-\infty}^{\infty} L(\theta, \phi(x)) f_{x|\theta} dx$$

El riesgo Bayesiano será la esperanza del riesgo con respecto a una función de densidad a priori de θ , dada por la forma

$$R(F_\theta, \phi) = E[\hat{R}(\theta, \phi)] = \int_{\Theta} \hat{R}(\theta, \phi) f_\theta(\theta) d\theta$$

$$R(F_\theta, \phi) = E[\hat{R}(\theta, \phi)] = \int_{\Theta} \int_X L(\theta, \phi(x)) f_{x,\theta}(x, \theta) dx d\theta$$

El estimador Bayesiano óptimo $\phi(x)$ del riesgo promedio está dado por:

$$\underset{\phi(x)}{\operatorname{argmín}} R(F_\theta, \phi(x)).$$

Según los tipos de costos, se obtienen los siguientes estimadores:

1. Dada una pérdida $L(\theta, \phi(x)) = (\theta - \phi(x))^2$, el estimador que minimiza el riesgo Bayesiano es la *Espezanza condicional*.

$$\phi(x) = \int_{\Theta} \theta \cdot f_{\theta|x}(\theta|X) d\theta.$$

2. Dada una pérdida de la forma $L(\theta, \phi(x)) = |\theta - \phi(x)|$ el estimador que minimiza el riesgo Bayesiano es la *Mediana*

$$\phi(x) = \operatorname{mediana} f_{\theta|x}(\theta|x).$$

3. Dada una pérdida uniforme de la forma

$$L(\theta, \phi(x)) = \begin{cases} 0 & \text{si } |\theta - \phi(x)| \leq \epsilon/2 \\ 1 & \text{si } |\theta - \phi(x)| > \epsilon/2 \end{cases}$$

El estimador que minimiza el riesgo bayesiano es la *Moda*. En el caso en que $\epsilon \rightarrow 0$ el estimador será el maximum a posteriori (o *MAP*):

$$\operatorname{argmáx}_{\theta} f_{\theta|x}(\theta|x). \tag{2.3}$$

2.3.2. La esperanza condicional

Como ya se ha definido, el estimador que minimiza el riesgo Bayesiano para el caso de un costo cuadrático $L(\theta, \phi(x)) = (\theta - \phi(x))^2$ es la esperanza condicional.

$$\hat{\theta} = \int_{\Theta} \theta f_{\theta|x}(\theta|x) d\theta \quad (2.4)$$

Ahora bien, tenemos que el riesgo Bayesiano esta dado por:

$$\begin{aligned} R(F_{\theta}, \phi(x)) &= E_{x,\theta}[L(\theta, \phi(x))] \\ &= E_{x,\theta}[(\theta - \phi(x))^2]. \end{aligned}$$

Notar que $R(F_{\theta}, \phi(x))$ es el error cuadrático medio de la distribución conjunta de θ y x , por lo que este estimador minimiza el Error Cuadrático Medio.

2.3.3. Estimación Paramétrica

Hasta ahora se ha considerado el parámetro como una variable aleatoria. Si ahora se considera que el parámetro θ es determinístico pero desconocido. El riesgo a minimizar viene dado por

$$R(F_{\theta}, \phi(x)) = \int_X (\theta - \hat{\theta})^2 f_{x|\theta}(x|\theta) dx.$$

Derivando e igualando a cero con respecto a θ , se llega a

$$\hat{\theta} = \theta.$$

Haciendo el procedimiento estándar para poder encontrar estimadores Bayesianos, no se logra obtener ningún resultado en el caso paramétrico, por lo que es necesario tener otro criterio en estos casos, como el principio de máxima verosimilitud.

Maximum-Likelihood o ML

Si la La función $f_{\theta|x}(\theta|x)$ de la ecuación (2.3) es diferenciable, entonces se cumple que:

$$\left. \frac{\partial \ln f_{\theta|x}(\theta|x)}{\partial \theta} \right|_{\theta=\hat{\theta}_{ML}} = 0$$

como

$$\begin{aligned}
f_{\theta|x}(\theta|x) &= \frac{f_{x|\theta} \cdot f_{\theta}}{f_x(x)} \\
\ln f_{\theta|x}(\theta|x) &= \ln f_{x|\theta}(x|\theta) + \ln f_{\theta}(\theta) - \ln f_x(x) \\
\left. \frac{\partial \ln f_{\theta|x}(\theta|x)}{\partial \theta} \right|_{\theta=\hat{\theta}} &= \left. \frac{\partial \ln f_{x|\theta}(x|\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}} + \left. \frac{\partial \ln f_{\theta}(\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}} = 0
\end{aligned}$$

Como θ en este caso no es una variable aleatoria, $f_{\theta}(\theta)$ se reduce a cero, lo que queda la siguiente maximización, que es precisamente el principio de *maximum likelihood* deducido de la regla de *maximum a posteriori*.

$$\left. \frac{\partial \ln f_{x|\theta}(x|\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}} = 0 \tag{2.5}$$

Es interesante recalcar que en este caso θ es un parámetro determinístico y desconocido. Este entra en contradicción con la función de densidad de probabilidad $f_{x|\theta}(x|\theta)$, por lo mismo, se habla más bien de *función de likelihood*, que cumple

$$f_{x|\theta}(x|\theta) = L(\theta, x) = \prod_{i=1}^m f_x(x_i|\theta)$$

en el caso i.i.d. la función conjunta y $L(\theta, x)$ coinciden.

Least Squares

Si bien el estimador *Least Squares* no se puede deducir directamente de la aproximación antes planteada, es importante nombrarla, dado que es la forma más habitual de estimar la posición del centroide de un objeto luminoso en la imagen.

Tal como su nombre lo indica, este estimador estima los parámetros de un sistema a través de la minimización del error cuadrado que existe entre las observaciones y el valor esperado de la función de densidad de probabilidad que está sujeto el problema.

Por lo tanto, sea $\vec{x} = (x_1, \dots, x_n)$ un vector de observación, y $E(x)$ la esperanza de las muestras dada una función de densidad de probabilidad en las que se distribuyen estas, entonces el estimador será

$$\hat{\theta} = \underset{\theta}{\operatorname{argmín}} \sum_{i=1}^n (x_i(\theta) - E(x_i(\theta)))^2 \tag{2.6}$$

Este estimador es ampliamente usado en debido a la facilidad computacional que implica implementarlo, pero en terminos estrictos, este estimador no cumple con ciertas garantías que pueden

dar otros estimadores, como puede ser *Maximum Likelihood* y sus propiedad de consistencia (convergencia al valor estimado), normalidad asintótica (su covarianza converge a la inversa de la matriz de fisher) y eficiencia (alcanza la Cota de Cramér-Rao con una alta cantidad de datos).

2.3.4. La Cota de Cramér-Rao

Es en muchas aplicaciones donde el cálculo explícito de la varianza del estimador es complejo de calcular, por lo que un análisis de los límites inferiores de estos estimadores puede ser una poderosa herramienta, en especial para conocer y comparar el desempeño entre estos, además conociendo el mínimo valor que pueden alcanzar los estimadores paramétricos, también es posible conocer cuán lejos están los resultados que se tienen con respecto al mejor caso posible. Existen muchas opciones para obtener una expresión de límites fundamentales para la mínima varianza, quizá la expresión más significativa viene dada por la cota de Cramér-Rao.

Sea $\hat{\theta}()$ un estimador insesgado de el parámetro determinístico θ . Si se define la función de *Likelihood* $L(I_1, \dots, I_n; \theta)$, cualquier estimador en función de n variables aleatorias de distribución $f_x(x; \theta)$ tiene como límite inferior

$$\text{Var}(\hat{\theta}) \geq \frac{1}{I_n(\theta)} \quad n: \text{número de mediciones} \quad (2.7)$$

donde

$$I(\theta) = \text{E} \left[\left(\frac{\partial \ln L(x; \theta)}{\partial \theta} \right)^2 \right] = -\text{E} \left[\frac{\partial^2 \ln L(x; \theta)}{\partial^2 \theta} \right] \quad (2.8)$$

se conoce como la información de Fisher del paramétrico θ , respecto a las mediciones.

Generalmente esta cota no es alcanzada por la varianza de los estimadores de θ , sin embargo, existe una condición necesaria y suficiente que garantiza la alcanzabilidad de este límite. Si es posible encontrar esta descomposición.

$$\frac{d \ln(I_1, \dots, I_n; \theta)}{d\theta} = A(\theta) \cdot (\hat{\theta}(I_1, \dots, I_n) - \theta) \quad (2.9)$$

entonces $\text{Var}(\hat{\theta}(I_1, \dots, I_n)) = 1/I_n(\theta)$ solo si $A(\theta)$ es una función exclusiva del parámetro θ (y particularmente no depende de las variables observadas I_i)

2.3.5. La Cota de Cramér-Rao Bayesiana (Desigualdad de Van Trees) [10]

Tal como en la versión paramétrica, el caso Bayesiano también incluye límites fundamentales con el cuál analizar o comparar el desempeño de sus estimadores. En este caso se tendrá que la

desigualdad dependerá tanto de las mediciones como también del espacio de parámetros que se esté considerando.

Sea θ una variable aleatoria y x el vector de observación. El Error Cuadrático Medio (MSE o *Mean Square Error*) de cualquier estimador satisface:

$$\mathbb{E}[(\hat{\theta} - \theta)^2] \geq \left(\mathbb{E} \left[\left(\frac{\partial \ln f_{x,\theta}(x, \theta)}{\partial \theta} \right)^2 \right] \right)^{-1} \quad (2.10)$$

$$= \left(-\mathbb{E} \left[\frac{\partial^2 \ln f_{x,\theta}(x, \theta)}{\partial^2 \theta} \right] \right)^{-1} \quad (2.11)$$

Dado que es la esperanza conjunta, se obtiene finalmente la siguiente desigualdad [3]

$$\mathbb{E}[(\hat{\theta} - \theta)^2] \geq \frac{1}{\mathbb{E}(I_n(\theta)) + I(\lambda)} \quad (2.12)$$

con

$$I(\lambda) = \mathbb{E} \left[\frac{\partial^2 \ln f_{x_c}(x_c)}{\partial x_c^2} \right] \quad (2.13)$$

donde $I(\theta)$ es la información de Fisher de $f_{x|\theta}(x|\theta)$ y $I(\lambda)$ es la información de Fisher de $f_\theta(\theta)$.

Adicionalmente y de manera análoga al caso paramétrico, la ecuación (2.12) se cumple con igualdad si

$$\frac{\partial f_{x,\theta}(x, \theta)}{\partial \theta} = k(\hat{\theta} - \theta). \quad (2.14)$$

Notar que en este caso, a diferencia que en el caso de la cota de Cramér-Rao clásico, k no puede depender del parámetro θ .

Si se deriva nuevamente la expresión (2.14) y expresa únicamente en términos de la densidad a posteriori se tiene que:

$$\frac{\partial^2 \ln f_{\theta|x}(\theta|x)}{\partial \theta^2} = -k. \quad (2.15)$$

Esto da una relación que debe cumplir la p.d.f (función de densidad de probabilidad por sus siglas en inglés) para que la cota de Cramér-Rao Bayesiana se cumpla con igualdad.

2.3.6. Comparación entre estimación paramétrica y Bayesiana

En estricto rigor, no es posible comparar la varianza de un estimador paramétrico con respecto al error cuadrático medio de un estimador Bayesiano, lo mismo pasa con sus límites fundamentales.

Por ello es necesario llevar a unidades comparables a la cota de Cramér-Rao para el caso paramétrico, realizando un promedio de $I_n(\theta)$ respecto a todos los valores posibles de θ . Dicho promedio es respecto a las estadísticas a priori, es decir.

$$MSE_{\sigma_{CR}} = E_{x_c \sim N(\theta, \mu)} \left(\frac{1}{I(\theta)} \right). \quad (2.16)$$

Este valor promedio se conoce como la *Esperanza de Cramér-Rao* (MCR por sus siglas en inglés). Notar que no depende del parámetro, más aún, dada la desigualdad de Jensen, se puede demostrar que:

$$\frac{1}{E(I(\theta))} \leq E_{x_c \sim N(\mu, \sigma)} \left(\frac{1}{I(\theta)} \right). \quad (2.17)$$

Finalmente, como $I(\theta) \geq 0$, entonces

$$\frac{1}{E(I(\theta)) + I(\lambda)} \leq E_{x_c \sim N(\mu, \sigma)} \left(\frac{1}{I(\theta)} \right), \quad (2.18)$$

en otras palabras:

$$\sigma_{BCR} \leq \sigma_{ECR} \quad (2.19)$$

De esta forma, en términos del error cuadrático medio, la cota Bayesiana de Cramér-Rao siempre será menor que la esperanza de Cramér-Rao. Este resultado trae consecuencias importantes, debido a que, en promedio, el tener información a priori del parámetro siempre mejorará la precisión que el sistema tendrá, noción que intuitivamente se tiene al utilizar este enfoque y que se ratifica en momento de tener este resultado. Ahora bien, al hablar en términos promedios, también existe otro elemento que también es importante de recalcar; el error ya no depende del parámetro, o en este caso, de la posición específica del pixel donde se está haciendo la medición. Esta es una situación destacable, dado que dentro de cada pixel es bastante complejo, sino imposible, saber el lugar exacto en donde el centroide de la estrella está realmente, por lo que el estudio entre estimadores y la cota no se puede hacer hasta después del cálculo de la varianza de la estimación (asumiendo de que se tratan de estimadores insesgados en el caso paramétrico). A diferencia de la Cota de Cramér-Rao, donde solamente se observa el parámetro punto por punto en el espacio del arreglo de pixeles, este análisis da un comportamiento promedio del sensor, que tiene una gran utilidad al analizar con un solo indicador su comportamiento global.

Capítulo 3

Implementación

3.1. Descripción del problema

La determinación de la posición del centroide de la estrella es una aplicación de las herramientas mostradas en el capítulo anterior. En esta memoria se revisará solamente este problema unidimensionalmente, con la intención de reducir el problema y obtener conclusiones que sirvan para analizar sistemas más complejos.

Supóngase una cámara CCD como un arreglo unidimensional de n pixeles, cada uno mide el flujo (cuentas) I_i que recibe de una fuente luminosa, con $i = 1, \dots, n$ mediciones independiente y no-identicamente distribuida. El valor en cada pixel es una realización aleatoria dada por una distribución de Poisson de valor esperado λ_i , Esta adquisición ruidosa de una estrella es dependiente de un modelo que infiere cómo debería comportarse la medición, quedando implícito en el valor λ_i .

Por propiedades de una realización de Poisson, su valor esperado $E(I_i)$ es igual al parámetro que la define, lo que significa que

$$E(I_i) = \lambda_i. \tag{3.1}$$

Es así que λ_i no solamente corresponde al modelo propuesto sino también a la esperanza, lo que le da la facultad de inferir parámetros del modelo a través de las mediciones.

Para efectos prácticos, el valor esperado λ_i en un arreglo de pixeles define “lo que se debería ver” como estrella si la medición no fuera un proceso aleatorio. La colección del flujo en este arreglo unidimensional correspondería a la medición del objeto discretizado más elementos externos que aumentan el valor basal del flujo por pixel.

En este modelo simplificado, la esperanza de la medición del flujo en el pixel depende únicamente de la posición de este, el cual contiene la información del centroide de la estrella. Formalmente se tiene que

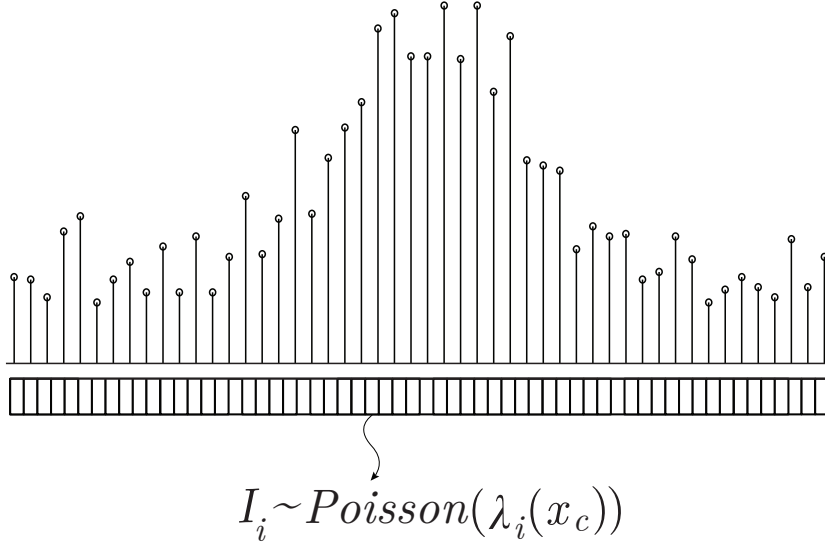


Figura 3.1: Medición del arreglo unidimensional de píxeles

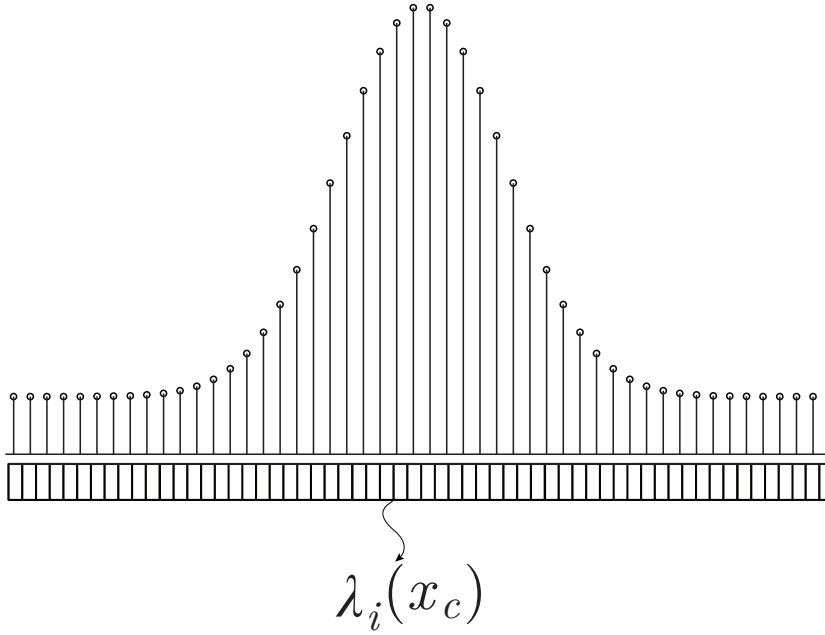


Figura 3.2: Discretización de una estrella en un arreglo unidimensionales de píxeles

$$\lambda_i(x_c) = \tilde{F}_i(x_c) + B_i. \tag{3.2}$$

El valor del flujo proviene de dos fuentes distintas; mientras que el aporte $\tilde{F}_i(x_c)$ es el valor integrado de la estrella en el i -ésimo píxel, que es la fuente de luz que se quiere analizar y extraer

características, B_i es el flujo que proviene tanto de elementos espúreos que reciben los pixeles del cielo (flujo de otras estrellas, rayos cósmicos, etc.) como también razones propias de la física del sensor, como lo son el ruido de lectura o *RON* (*Read-out Noise*) y el ruido producido por electrones térmicamente excitados, llamado corriente oscura o *Dark Current*. Estos valores no dependen directamente del parámetro medido y son propios de cada pixel, pero como su variación es pequeña [6] entonces

$$B_i \approx B \quad (3.3)$$

donde

$$B = f_s \cdot \Delta x + \frac{D + \text{RON}^2}{G} [\text{ADU}]. \quad (3.4)$$

Y donde f_s es el aporte del cielo en el ruido, el cual está ponderado por largo de cada pixel (medido en [arcsec]), mientras que D y RON son la corriente oscura y el ruido de lectura del detector medido por pixel respectivamente, en unidades de electrones (e^-). Este resultado es interesante, ya que se puede apreciar que para aplicaciones astronómicas clásicas, el ruido del cielo determinará el comportamiento global del ruido de fondo, esto es porque f_s depende directamente del tiempo y para exposiciones prologadas su valor crece, haciendo que el ruido de lectura y de corriente oscura tengan una menor incidencia en la forma del ruido total.

Con respecto al aporte de la estrella en el flujo I_i , este dependerá también de la *Point Spread Function* o PSF $\phi(x)$ (función que define la forma de la objeto medido) y del pixel donde se está recibiendo la información, por lo tanto

$$\tilde{F}_i = \tilde{F} \cdot g_i(x_c) \quad (3.5)$$

donde \tilde{F} es el flujo total y $g_i(x_c)$ la integración de la forma de la estrella en el pixel analizado.

En la Figura 3.3 se puede apreciar gráficamente como obtiene el valor del flujo del objeto luminoso por pixel, donde se hace un conteo de los fotones que han incidido en cada uno de los “huecos” del arreglo unidimensional. Cuantos fotones incidan en un segmento particular del arreglo vendrá dado por cuan puntual es la fuente de emisión y por tanto de $g_i(x_c)$, que está definido como:

$$g_i(x_c) = \int_{x_i - \frac{\Delta x}{2}}^{x_i + \frac{\Delta x}{2}} \phi(x, x_c) dx \quad (3.6)$$

donde x_i denota la posición central del pixel $i \in \{1, \dots, n\}$ y Δx el tamaño de pixel. En este modelo se asume que existe una cantidad suficiente de pixeles para recoger todo el valor del flujo, que se expresa en

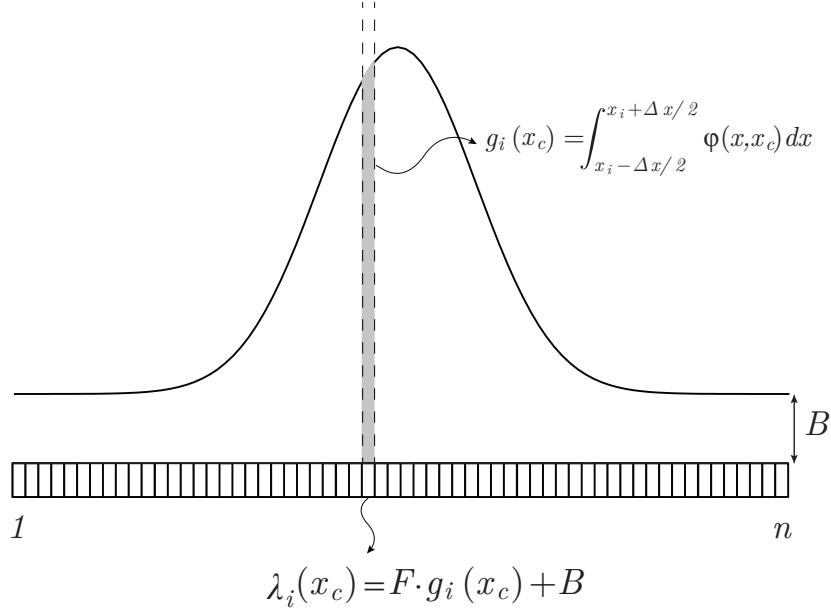


Figura 3.3: Funcionamiento simplificado de un arreglo de pixeles como un integrador de fotones

$$\sum_{i=1}^n \tilde{F}_i(x_c, \tilde{F}) = \tilde{F} \sum_{i=1}^n \tilde{F} \cdot g_i(x_c) = \tilde{F} \sum_{i=1}^n \tilde{F} \cdot \int_{x_i - \frac{\Delta x}{2}}^{x_i + \frac{\Delta x}{2}} \phi(x, x_c) dx = \tilde{F} \sum_{i=1}^n \tilde{F} \cdot \int_{-\infty}^{\infty} \phi(x) dx = \tilde{F} \quad (3.7)$$

esto último es porque la PSF es un valor normalizado, por lo que $\int_{-\infty}^{\infty} \phi(x) dx = 1$

3.1.1. Astrometría en CR y BCR

De Mendez et al. [6] se desprende que la cota de CR tiene la siguiente forma cerrada:

$$\sigma_{CR}^2 = 2\pi\sigma^2 \cdot \frac{B}{GF^2} \cdot \frac{1}{\sum_{i=1}^n \frac{(e^{-\gamma(x_i^-)} - e^{-\gamma(x_i^+)})^2}{1 + \frac{1}{\sqrt{2}\phi\sigma} \frac{F}{B} \int_{x_i - \frac{\Delta x}{2}}^{x_i + \frac{\Delta x}{2}} e^{-\gamma(x)} dx}}$$

donde es x_i la posición del centro del pixel i , Δx su tamaño y por lo tanto $x_i^- = x_i - \frac{\Delta x}{2}$ y $x_i^+ = x_i + \frac{\Delta x}{2}$ su posiciones inicial y final respectivamente. El funcional $\gamma(x)$ corresponde a

$$\gamma(x) = \frac{(x - x_c)^2}{2\sigma^2} \quad (3.8)$$

por lo tanto, se tiene que la información de Fisher está dada por

$$I(x_c) = \frac{GF^2}{2\pi\sigma^2 B} \sum_{i=1}^n \frac{(e^{-\gamma(x_i^-)} - e^{-\gamma(x_i^+)})^2}{1 + \frac{1}{\sqrt{2}\phi\sigma} \frac{F}{B} \int_{x_i - \frac{\Delta x}{2}}^{x_i + \frac{\Delta x}{2}} e^{-\gamma(x)} dx} \quad (3.9)$$

Utilizando la desigualdad de Van Trees de la ecuación (2.12) y reemplazando la información de Fisher, se obtiene:

$$\mathbb{E}[(\hat{\theta} - \theta)^2] \geq \left(\mathbb{E}_{x_c \sim N(\theta, \mu)} \left(\frac{GF^2}{2\pi\sigma^2 B} \sum_{i=1}^n \frac{(e^{-\gamma(x_i^-)} - e^{-\gamma(x_i^+)})^2}{1 + \frac{1}{\sqrt{2}\phi\sigma} \frac{F}{B} \int_{x_i - \frac{\Delta x}{2}}^{x_i + \frac{\Delta x}{2}} e^{-\gamma(x)} dx} \right) + \mathbb{E}_{x_c \sim N(\theta, \mu)} \left(\frac{\partial^2 \ln f_{x_c}(x_c)}{\partial x_c^2} \right) \right)^{-1}. \quad (3.10)$$

Si bien en este caso se ha dejado explícito el valor de la información de Fisher, para poder caracterizar la cota de Cramér-Rao Bayesiana en astrometría es necesario dejar explícita la función de densidad de probabilidad a priori que se usará al estimar el parámetro. Este funcional determinará el resultado tanto de la esperanza de la información de Fisher como también el resultado final de la información de Fisher a priori. Para efectos de esta memoria, se analizarán los siguientes casos:

Distribución a priori Uniforme

Es fácil ver que la información de Fisher a priori en este caso es

$$I(\lambda) = 0,$$

por lo que la desigualdad es muy parecida a Cramér-Rao

$$\mathbb{E}[(\hat{x}_c - x_c)^2] \geq \frac{1}{\mathbb{E}(I(x_c))} \quad (3.11)$$

Este sería el caso equivalente al enfoque paramétrico, dado que la información a priori toma como supuesto que cada posición en el arreglo de pixeles tiene igual probabilidad de contener al centroide de la PSF, por lo que en teoría ambos resultados deberían ser similares en desempeño. Sin embargo, por la desigualdad de Jensen, el error cuadrático medio del caso Bayesiano será siempre menor al paramétrico en término de sus límites fundamentales.

Distribución a priori Normal $N(x_c, \sigma_{priori})$

La densidad de probabilidad a priori asociada a la estimación será la siguiente:

$$\begin{aligned}
f_{x_c}(x_c) &= \frac{1}{\sqrt{2\pi}\sigma_{priori}} e^{-\frac{(x_c-\mu)^2}{2\sigma_{priori}^2}} \\
\ln f_{x_c}(x_c) &= -\ln(\sqrt{2\pi}\sigma_{priori}) - \frac{(x_c - \mu)^2}{2\sigma_{priori}^2} \\
\frac{\partial^2 \ln f_{x_c}(x_c)}{\partial x_c^2} &= -\frac{1}{\sigma_{priori}^2}
\end{aligned}$$

por lo que la información de Fisher a priori es:

$$I(\lambda) = E_{x_c \sim N(\theta, \mu)} \left[\frac{\partial^2 \ln f_{x_c}(x_c)}{\partial x_c^2} \right] = \frac{1}{\sigma_{priori}^2} \quad (3.12)$$

En una forma simplificada, la desigualdad de Cramér-Rao Bayesiana queda de la siguiente forma

$$E[(\hat{x}_c - x_c)^2] \geq \frac{1}{E(I(x_c)) + \frac{1}{\sigma_{priori}^2}} \quad (3.13)$$

Este resultado muestra que la única variable relevante que se agrega a la ecuación es la desviación estandar de la p.d.f a priori. Se incorpora entonces el “error ” previo a la estimación. Dada la forma de la ecuación, se espera que la cota Bayesiana siempre sea menor o igual a este valor, además, dado el resultado de la desigualdad de Jensen, también será menor que la Cota de Cramér-Rao paramétrico.

3.2. Estimador Bayesiano del Problema

Tal cual se ha comentado en capítulos anteriores, el estimador Bayesiano óptimo es la esperanza condicional. Este será ocupado para el análisis en este reporte por ser un estimador suficiente y que alcanza la mínima varianza, condiciones que aseguran mostrar el mejor caso alcanzable al utilizar este tipo de enfoque al estimar centroides de estrellas.

3.2.1. Caso Astrométrico

La esperanza condicional, tal como se ha mostrado anteriormente, tiene la siguiente forma

$$\hat{x}_c = E(x_c|x)$$

la esperanza está dada por la función de probabilidad del parámetro dada las observaciones, escrito en su forma integral se puede descomponer de la siguiente forma

$$\hat{x}_c = \int x_c f_{x_c|x}(x_c|x) dx_c,$$

usando el teorema de Bayes en la p.d.f.

$$\hat{x}_c = \int x_c \cdot \frac{f_{x|x_c} \cdot f_{x_c}}{f_x} dx_c.$$

Como f_x no depende del parámetro, se puede independizar de la integral, además por teorema de probabilidades totales, $f_x = \int f_{x|x_c} \cdot f_{x_c} dx_c$. La expresión final del estimador es la siguiente

$$\hat{x}_c = \frac{\int x_c \cdot f_{x|x_c} \cdot f_{x_c} dx_c}{\int f_{x|x_c} \cdot f_{x_c} dx_c}, \quad (3.14)$$

por lo tanto, tanto el denominador como el numerador se pueden escribir independientemente. Como también comparten el término $f_{x|x_c} \cdot f_{x_c}$, será este el que se calculará primero para después proceder a los cálculos de las integrales.

Nuevamente, dependiendo de la información a priori del sistema, se tendrá un estimador. Los casos a estudiar nuevamente serán una p.d.f a priori uniforme y otra normal.

Información a priori $N(\mu, \sigma)$

Con $f_{x_c} = N(\mu, \sigma)$ y $f_{x|x_c}$ una distribución conjunta de n realizaciones de Poisson i.n.d. (independientes y no-identicamente distribuidas). Se calculará $f_{x|x_c} \cdot f_{x_c}$ para luego calcular las integrales correspondientes.

$$\begin{aligned} f_{x|x_c} &= \prod_{i=1}^n f_{x_i|x_c}(x_i|x_c) \\ &= \prod_{i=1}^n \frac{e^{-\lambda_i(x_c)}}{x_i!} \lambda_i^{x_i} \\ \Rightarrow f_{x|x_c} \cdot f_{x_c} &= \left(\prod_{i=1}^n \frac{e^{-\lambda_i(x_c)}}{x_i!} \lambda_i^{x_i} \right) \cdot N(x_c, \sigma) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_c-\mu)^2}{2\sigma^2}} \cdot \left(\prod_{i=1}^n \frac{e^{-\lambda_i(x_c)}}{x_i!} \lambda_i^{x_i} \right). \end{aligned}$$

Las integrales entonces serán

$$\int f_{x|x_c} \cdot f_{x_c} dx_c = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{x_c} e^{-\frac{(x_c-\mu)^2}{2\sigma^2}} \cdot \left(\prod_{i=1}^n \frac{e^{-\lambda_i(x_c)}}{x_i!} \lambda_i^{x_i} \right) dx_c \quad (3.15)$$

y

$$\int x_c \cdot f_{x|x_c} \cdot f_{x_c} dx_c = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{x_c} x_c \cdot e^{-\frac{(x_c-\mu)^2}{2\sigma^2}} \cdot \left(\prod_{i=1}^n \frac{e^{-\lambda_i(x_c)}}{x_i!} \lambda_i^{x_i} \right) dx_c. \quad (3.16)$$

Con las Ecuaciones (3.15) y (3.16), obtenemos finalmente las expresiones del numerador de denominador de la esperanza condicional (Ecuación (3.14)).

Información a priori $U(Lim_{Inf}, Lim_{Sup})$

Supongamos el caso de la información uniforme de rango $[Lim_{Inf}, Lim_{Sup}]$, como tal cual ya se presentó en la Sección 3.2.1 , basta con calcular

$$\hat{x}_c = \frac{\int x_c \cdot f_{x|x_c} \cdot f_{x_c} dx_c}{\int f_{x|x_c} \cdot f_{x_c} dx_c}.$$

Como $f_{x_c}(x_c) = \frac{1}{Lim_{Sup} - Lim_{Inf}}$ el estimador óptimo es el siguiente

$$\hat{x}_c = \frac{\int_{x_c} x_c \cdot \left(\prod_{i=1}^n \frac{e^{-\lambda_i(x_c)} \lambda_i^{x_i}}{x_i!} \right)}{\int_{x_c} \left(\prod_{i=1}^n \frac{e^{-\lambda_i(x_c)} \lambda_i^{x_i}}{x_i!} \right)} \quad (3.17)$$

3.3. Implementación de la Esperanza Condicional en Astrometría

Para la implementación computacional del estimador, es necesario descomponer la expresión (3.14) de forma eficiente, como se darán a conocer en los siguientes pasos

1. Generar una matriz de datos de λ_{i,x_c} : Dado que esta expresión depende tanto de la posición del centroide de la PSF como también del valor de esta en cada pixel, el resultado tendrá una matriz que contiene por fila contiene los pixeles del arreglo de CCD asociados a una PSF con un centroide dado, mientras que por columna se varía la posición de este en el arreglo de pixeles.

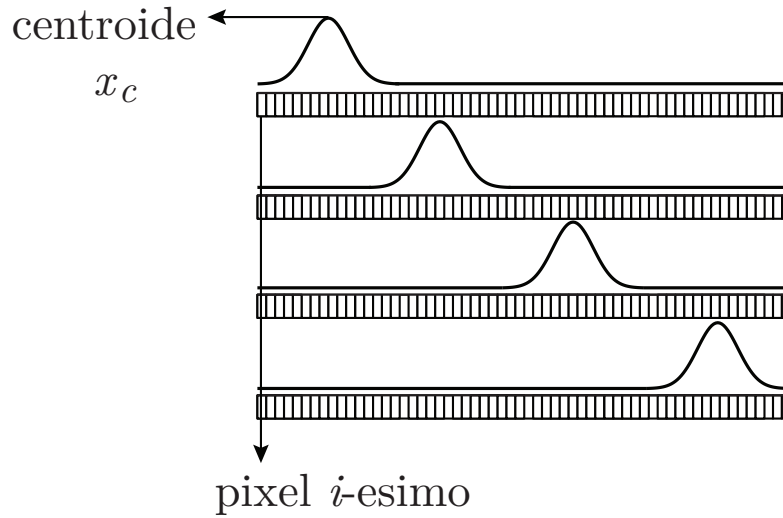


Figura 3.4: Matriz de datos de λ_{i,x_c} , donde cada fila completa contiene a una estrella de acuerdo al modelo planteado, mientras que por cada columna se encuentra la posición de esta, que es el parámetro a estimar

2. Calcular $f_{\vec{n}|x_c} = \prod_{i=1}^n \frac{e^{-\lambda_i(x_c)}}{x_i!} \lambda_i^{x_i}$ y multiplicar por $f_{x_c} = N(x_c, \sigma_{priori})$: Dado los órdenes de magnitud de las observaciones y de λ_{i,x_c} (superiores a 100 [ADU]) existe un potencial problema de *overflow* (los números sean más grandes que lo que se pueda almacenar). Una alternativa es resolver el problema logarítmicamente:

$$\begin{aligned}
f_{\vec{n}|x_c} &= \prod_{i=1}^n \frac{e^{-\lambda_i(x_c)}}{x_i!} \lambda_i^{x_i} \\
\ln(f_{\vec{n}|x_c}) &= \ln\left(\prod_{i=1}^n \frac{e^{-\lambda_i(x_c)}}{x_i!} \lambda_i^{x_i}\right) \\
&= \sum_{i=1}^n \ln\left(\frac{e^{-\lambda_i(x_c)}}{x_i!} \lambda_i^{x_i}\right) \\
&= \sum_{i=1}^n -\lambda_i(x_c) - \ln(x_i!) + x_i * \ln(\lambda_i^{x_i}).
\end{aligned}$$

El término $\ln(x_i!)$ se puede escribir como

$$\begin{aligned}
\ln(x_i!) &= \ln\left(\prod_{k=1}^{x_i} k\right) \\
&= \sum_{k=1}^{x_i} \ln(k)
\end{aligned} \tag{3.18}$$

lo que finalmente reduce el cálculo a

$$f_{\vec{n}|x_c} = \sum_{i=1}^n -\lambda_i(x_c) - \sum_{k=1}^{x_i} \ln(k) + x_i \cdot \ln(\lambda_i^{x_i})$$

3. Obtener un vector producto entre la función de densidad de probabilidad (p.d.f.) de los datos dado el parámetro, multiplicado por la p.d.f a priori, esta se le llamará f_{xtheta} (nombre que tendrá en la implementación computacional)

$$f_{\text{xtheta}} = f_{\vec{n}|x_c} \cdot N(\mu, \sigma_{\text{priori}})$$

este valor dependerá todavía del parámetro (el centroide de la PSF).

4. Calcular f_x como

$$f_x = \int f_{\text{xtheta}} dx_c,$$

donde se usa la integración numérica por medio de trapezoides.

5. Determinar el estimador de la siguiente forma

$$\hat{x}_c = \int x_c \cdot \frac{f_{x\theta}}{f_x} dx_c. \quad (3.19)$$

Nuevamente se procede con el método de trapezoide para integrar.

3.4. Implementación de la Varianza Condicional en Astrometría

Para medir la precisión del estimador Esperanza Condicional, es necesario calcular su error cuadrático medio. La ecuación a desarrollar para encontrar este valor es la siguiente

$$\text{MSE} = \int_X \int_{X_C} (x_c - \hat{x}_c)^2 f_{x,x_c} dx dx_c \quad (3.20)$$

Dado que $f_{x,x_c} = f_{x|x_c} \cdot f_{x_c}$ la integral y por lo tanto, su cómputo numérico se pueden separar en dos términos:

$$\text{MSE} = \int_{X_C} \int_X (x_c - \hat{x}_c)^2 f_{x,x_c} dx dx_c \quad (3.21)$$

$$= \int_{X_C} \left(\int_X (x_c - \hat{x}_c)^2 f_{x|x_c} dx \right) f_{x_c} dx_c. \quad (3.22)$$

$$(3.23)$$

Si se utiliza la Ley de los Grandes Números, para una cantidad de muestras apropiadas de datos de pixeles, se tiene que su varianza es

$$\text{MSE} = \int_{X_C} \left(\frac{1}{N} \sum_{i=1}^N (x_c - \hat{x}_c)^2 \right) f_{x_c} dx_c, \quad (3.24)$$

donde N será el número de iteraciones necesarias para que la aproximación sea correcta. Para el caso de la varianza que depende de la función de densidad de probabilidad f_{x_c} también se puede usar la Ley de los Grandes Números, obteniéndose como resultado

$$\text{MSE} = \frac{1}{M} \sum_{j=1}^M \left(\frac{1}{N} \sum_{i=1}^N (x_c - \hat{x}_c)^2 \right) \quad (3.25)$$

dado que los argumentos no dependen del índice, se puede realizar una gran sumatoria

$$\text{MSE} = \frac{1}{P} \sum_{k=1}^P (x_c - \hat{x}_c)^2, \quad (3.26)$$

con $P = M \cdot N$, este valor determinará la cantidad de iteraciones necesarias para que el resultado converja a la solución real. Como se puede ver, la cantidad será grande debido a que se está aproximando una doble integral.

En resumidas cuentas, la implementación se reduce a lo siguiente:

1. Se obtiene un valor de de centroide x_c a través de una realización aleatoria definido por la función de densidad de probabilidad a priori $N(\mu, \sigma_{priori})$. En este caso, se tomará que el

centro de la normal de la p.d.f. a priori está en la posición real del valor de la estrella, por lo que

$$\mu_{priori} = x_c.$$

Posteriormente se realiza una simulación de una medición ruidosa de una estrella de acuerdo a la distribución de Poisson de parámetro $\lambda_i(x_c)$ en la posición obtenida por la realización aleatoria anterior, de esta manera se tendrá como realización una medición condicionada a x_c , que corresponderá a $f_{vecx|x_c}$

De esta manera, se aproximan las p.d.f. a priori f_{x_c} con una realización aleatoria de acuerdo a la normal $N(\mu_{priori}, \sigma_{priori})$ y se obtiene la realización aleatoria de la adquisición de datos condicionada a esa posición, siendo esta por lo tanto una realización aleatoria de

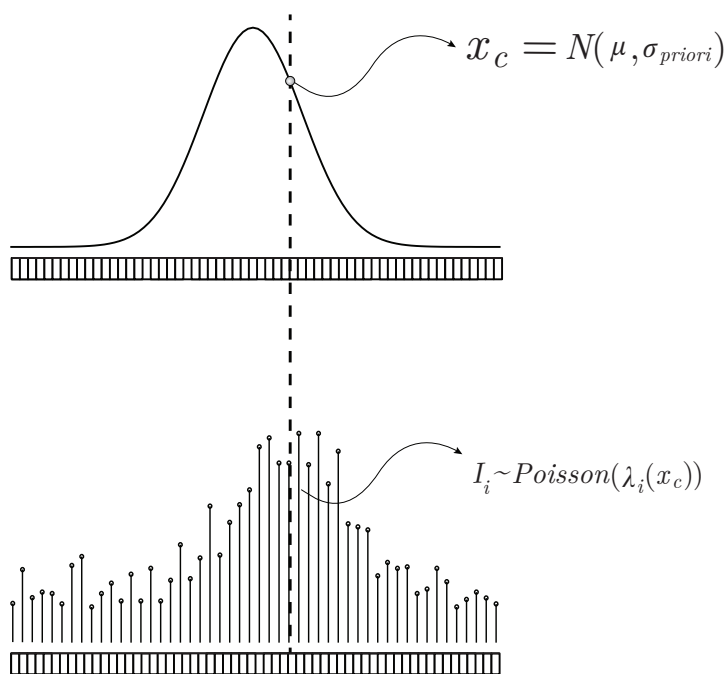


Figura 3.5: Creación de una realización en virtud de la variable aleatoria de x_c

2. Se obtiene el estimador \hat{x}_c
3. Se calcula la varianza para la realización i-ésima, $Var_i = (\hat{x}_c - x_c)^2$.
4. Se guarda el valor obtenido y repite el proceso para otro valor aleatorio x_c definido por la densidad de probabilidad a priori.

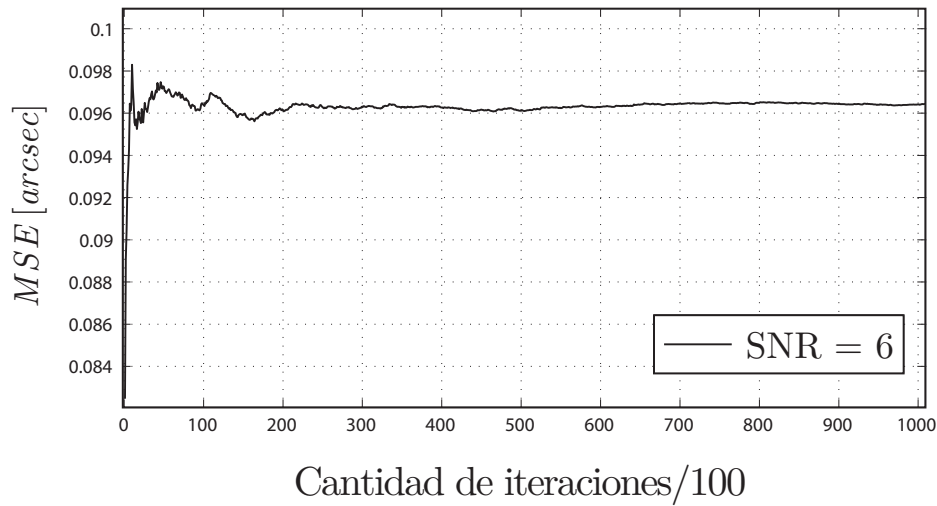


Figura 3.6: MSE acumulado en función de la cantidad de iteraciones, para un $FWHM = 1''$, $SNR = 6$, $fs = 2000[ADU/arcsec]$, $D = 0$, $RON = 5 [e^-]$

Ya que en este proceso lo que se está haciendo es obtener a través de la Ley de los Grandes Números el valor aproximado del error, es necesario tener una gran cantidad de muestras. A través de un análisis experimental se pudo deducir que la cantidad óptima de iteraciones, para el peor caso en hipótesis (baja relación señal-ruido), se necesitaron alrededor de treinta mil iteraciones para lograr un resultado óptimo en términos del error de aproximación de la varianza condicional. La Figura 3.6 muestra el MSE acumulado en función de la cantidad de iteraciones, donde se puede ver el valor de estabilización a cien mil iteraciones.

Capítulo 4

Análisis

En esta sección se presentará un análisis cuantitativo de los límites fundamentales del proyecto y su comparación con el estimador Bayesiano óptimo, la Esperanza Condicional. Se medirá su desempeño en función del error cuadrático medio, el cual se analizará como cambia al variar su valor para distintos tamaños de pixel y sus ganancias relativas con respecto al enfoque clásico paramétrico y con respecto a error a priori.

Finalmente se hace un análisis numérico de los distintos estimadores para finalmente analizar cuán cercana está la Esperanza Condicional de la Cota de Cramér-Rao Bayesiana.

4.1. Análisis Numérico

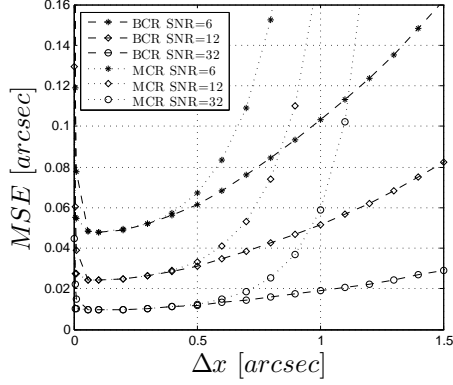
4.1.1. MSE con respecto al tamaño del pixel

Para este caso, solamente se tendrá en consideración el caso de una p.d.f. a priori normal, ya que a diferencia del caso uniforme, existe una variable que controlar y modificar, que es la desviación estandar a priori (σ_{priori}).

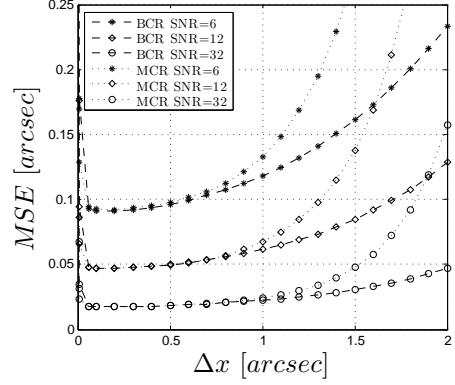
Revisando la Ecuación (3.13), se puede esperar que las diferencias entre Cramér-Rao Bayesiano y la Esperanza de Cramér-Rao ocurran cuando $\sigma_{priori} \ll 1$, caso en que $I(\lambda)$ es comparable al de $E(I_n(x_c))$.

Con respecto al tamaño del pixel, un aumento en su valor repercute rápidamente en un aumento en el error. El *MSE* del caso paramétrico crece significativamente para valores mayores a $0.6[\text{arcsec}]^1$, sin embargo, este aumento es atenuado en el caso Bayesiano, ya que al disminuir el valor de $E(I_n(x_c))$, la fuente de los información $I(\lambda)$ que proveniente de la p.d.f. a priori se mantiene constante, generando un resultado del error que tiende asintóticamente σ_{priori}^2 , como se verá adelante. Estas tendencias se muestran para todos los rangos de señal-ruido estudiados.

¹Un arco de segundo (o *arcsec*) es una medida astronómica de cuando mide un objeto angularmente en el cielo o bóveda celeste. El error se puede medir en medidas centradas en el sensor (la cámara CCD en este caso a través de una medida de distancia entre pixeles) o en el cielo, está última referencia es la de utilidad astronómica. Para esta



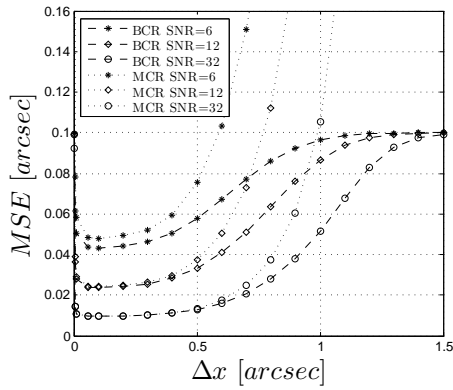
(a) $FWHM = 0.5''$



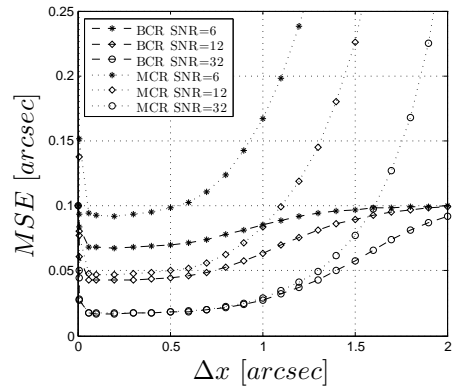
(b) $FWHM = 1''$

Figura 4.1: Error cuadrático medio en función del tamaño del pixel, para distintos $FWHM$. En ambos casos la información a priori es normal de $\sigma_{priori} = 0.5''$. La línea punteada corresponde a valores de BCR y la línea entrecortada a valores de MCR.

Como se puede ver en la Figura 4.1, el valor de $FWHM$ determina cuan rápido aumenta el error cuadrático medio en función del tamaño del pixel, siendo cada vez menor al hacer la estrella simulada más ancha. Esto es intuitivo, ya que al ser más pequeña la estrella, más difícil será muestrear la imagen y conocer su forma real, cosa que se atenúa en el caso de tener estrellas con un $FWHM$ mayor.



(a) $FWHM = 0.5''$



(b) $FWHM = 1''$

Figura 4.2: Error cuadrático medio en función del tamaño del pixel, para distintos $FWHM$. En ambos casos la información a priori es normal de $\sigma_{priori} = 0.1''$. La línea punteada corresponde a valores de BCR y la línea entrecortada a valores de MCR.

El aumento del valor del $FWHM$ se traduce en una mayor ganancia del enfoque Bayesiano. Para el caso de un $FWHM$ pequeño existe poca diferencia en el rango de tamaños de pixel de interés (cerca de $0.2[arcsec]$), donde la estrella está bien caracterizada a través de un buen muestreo de memoria, se utilizó un pixel de tamaño unitario y una relación $arcsec/tamaño\ pixel$ igual a $0,2 [arcsec/tamaño\ pixel]$

señal recibida.

Al cambiar el tamaño del $FWHM$ y de Δx , el muestreo pierde completamente su valor de información ($E(I_n(x_c))$ se hace pequeño), teniendo como único valor relevante lo que pueda dar la p.d.f a priori, estabilizándose en el valor de σ_{priori} , esto se puede ver en la Figura 4.3

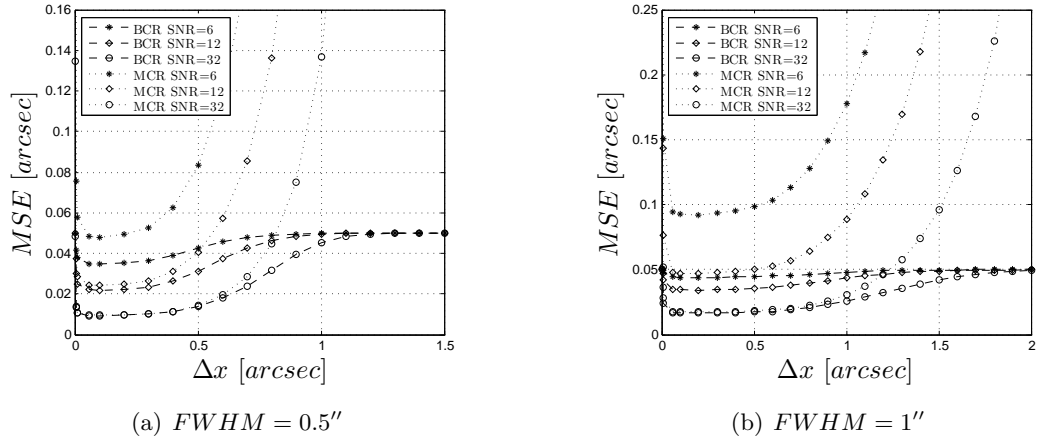


Figura 4.3: Error cuadrático medio en función del tamaño del pixel, para distintos $FWHM$. En ambos casos la información a priori es normal de $\sigma_{priori} = 0.05''$. La línea punteada corresponde a valores de BCR y la línea entrecortada a valores de MCR.

Al aumentar la información de Fisher a priori, usando un σ_{priori} de $0.05 [arcsec]$, se observan las mismas tendencias que para un σ_{priori} de $0.1 [arcsec]$. Las diferencias entre ambas curvas son significativas en los casos de una relación de señal-ruido bajas, llegando incluso a ordenes de magnitud de diferencia entre el enfoque Bayesiano y no Bayesiano (BCR llega a ser solo el 20% del valor de MCR en el caso de una relación de señal ruido de 6 y un tamaño de pixel de $0.2 [arcsec]$). Estas ganancias se van perdiendo a medida que la relación de señal-ruido aumentan. Tal como en el caso anterior, la estrella es mejor muestreada y por lo tanto los datos aportan cada vez más información en relación a la información previa que se tenía de la estimación.

4.1.2. Ganancia Relativa del caso bayesiano y aporte de la distribución a priori

Como se ha mostrado previamente, es con un $FWHM$ alto donde se tendrán mayores diferencias entre BCR y MCR, por lo que se toman valores de $FWHM = 1''$ en este análisis.

Se puede observar también como existe un compromiso entre la ganancia total de usar enfoques Bayesianos y sesgar completamente el resultado a la información a priori que se tenga del sistema. Por ejemplo, en la Figura 4.4, para un σ_{priori} de $0.2 [arcsec]$, ganancia relativa para un señal-ruido de 6, es de un 10% y su ganancia con respecto a σ_{priori} es de un 60%, pero para un σ_{priori} de $0.2 [arcsec]$, se tendrá un 26% y 30% respectivamente, por lo que la tendencia es clara a sesgar el

resultado previamente al aumentar el σ_{priori} de la p.d.f a priori.

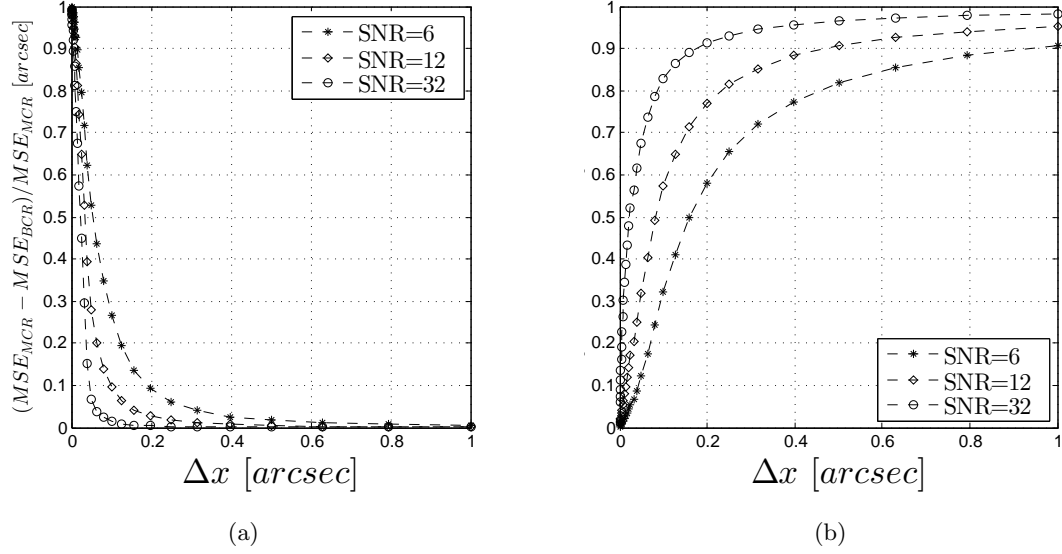


Figura 4.4: Ganancia relativa entre BCR y MCR, en función de σ_{priori} y el otro en la ganancia relativa de BCR con respecto al σ_{priori} en función del σ_{priori} , para $\Delta x = 0.2 [arcsec]$ y $FWHM = 1''$.

Como se ha mostrado previamente, es con un $FWHM$ alto donde se tendrán mayores diferencias entre BCR y MCR, por lo que se toman valores de $FWHM = 1''$ en este análisis.

Esta tendencia se repite para distintos valores de Δx , como se puede ver en 4.5, donde para $\Delta x = 0.7 [arcsec]$, donde se pueden tener ganancias relativas mayores sin necesariamente alterar en mayor medida la ganancia con respecto a σ_{priori} inicial.

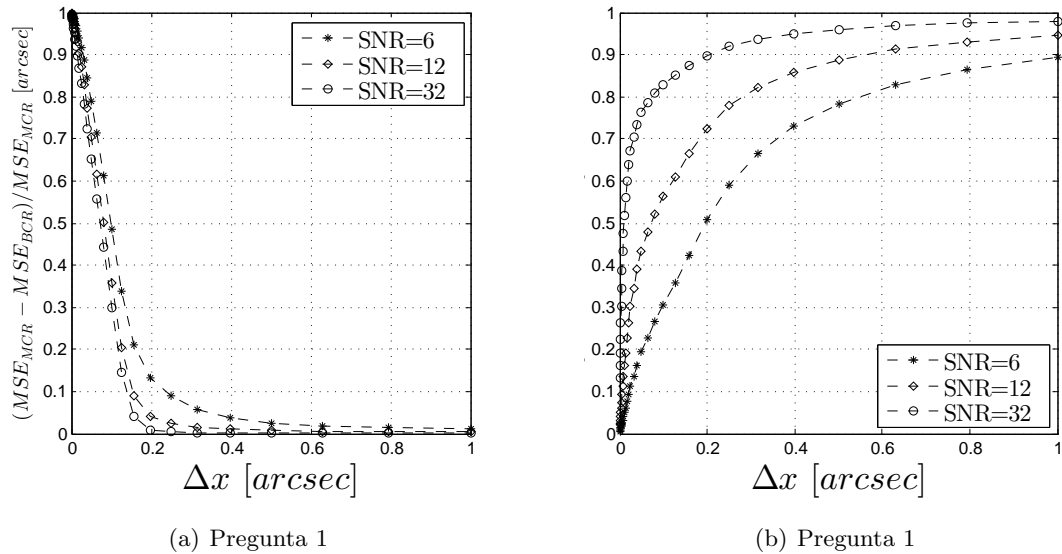


Figura 4.5: Ganancia relativa entre BCR y MCR, en función de σ_{priori} y el otro en la ganancia relativa de BCR con respecto al σ_{priori} en función del σ_{priori} , para $\Delta x = 0.7 [arcsec]$ y $FWHM = 1''$.

Esta tendencia se repite para distintos valores de Δx , como se puede ver en 4.5, donde para $\Delta x = 0.7 [arcsec]$ se pueden tener ganancias relativas mayores sin necesariamente alterar en mayor medida la ganancia con respecto a σ_{priori} inicial.

4.1.3. Brillo equivalente de los objetos

La pregunta es cuán certera debe ser la información previa del sistema, para, que dado un valor de flujo en el caso bayesiano, se tenga un resultado equivalente a un caso paramétrico. En términos matemáticos, la idea es resolver la siguiente identidad:

$$BFI(F, \sigma_{priori})^{-1} = MCR(F_{par}), \quad (4.1)$$

donde BFI son la siglas en inglés de la información de Fisher Bayesiano y MCR corresponde a las Esperanza de Cramér-Rao. F_{par} corresponde al flujo aparente de un caso equivalente sin información a priori, mientras que F correspondería al flujo real. En otras palabras, para un objeto de intensidad de flujo F se quiere encontrar el flujo F_{par} de un caso paramétrico (donde no se tienen una distribución a priori), tal que la Esperanza de Cramér-Rao y Cramér-Rao Bayesiano sean iguales.

La Figura 4.6 muestra el comportamiento de $\frac{F}{F_{par}} \in (0, 1)$ para distintos valores de σ_{prior} (con información de fisher de $I(\lambda) = 1/\sigma_{prior}$), para distintos flujos reales iniciales. En la Figura 4.6.a $F \in \{5000, 10000, 50000\}$ y en la Figura 4.6.b $F \in \{200, 500, 1000\}$. Como es esperado, se puede apreciar que para el caso en que $\sigma_{priori} \rightarrow \infty$, no existe aporte sustancial de la información a priori en el MSE de la cota de Cramér-Rao Bayesiana y por tanto, su valor estará dado solamente por la variación del flujo real. En este caso el flujo aparente F_{par} y F tendrán aproximadamente el mismo valor y se tendrá que $\frac{F}{F_{par}} \approx 1$. Por otro lado, al disminuir significativamente el valor de σ_{prior} , la información a priori de la Cota de Cramér-Rao Bayesiano aumenta y por tanto, la diferencia entre F y F_{par} , teniendo $\frac{F}{F_{par}} \rightarrow 0$. Ambos casos asintóticos ocurren para distintos casos de flujos reales, como se puede ver en ambas figuras.

Si bien en los casos en que el flujo real es alto los valores de σ_{priori} son poco realistas (ya que el error a priori es prácticamente conocer con total certeza la posición del centroide), los casos en que el flujo real tiene menor valor, se puede ver que puede haber un aporte del σ_{priori} , teniendo que para el caso de un $F = 200 [ADU]$ el flujo aparente aumenta en un 60 % con σ_{priori} de 0.1 [arcsec], lo que es un aumento significativo en términos porcentuales. Esto se puede entender como que, para alcanzar el mismo valor de error, en un caso paramétrico se necesita un 60 % más de flujo que podría no haberse necesitado si se hubiera incorporado en compensación información a priori de

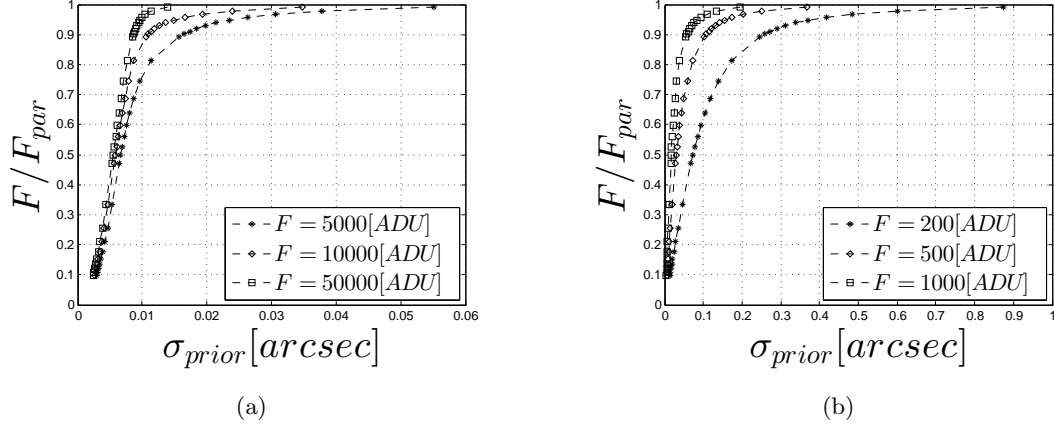


Figura 4.6: Brillo aparente de objetos, en relación a $\frac{F}{F_{par}}$ con respecto a σ_{prior} .

$\sigma_{prior} = 0.1$ [arcsec]. Esta tendencia se puede ver en menor o mayor medida para todo el rango de valores estudiados.

Para realizar esta curva, se mantuvieron los valores de *Dark Current* ($D = 0$ [e^-]), *Read-Out Noise* ($RON = 5$ [e^-]), *Sky Background* ($f_s = 2000$ [ADUs/arcsec]) y tamaño de pixel ($\Delta x = 0.2$ [arcsec]) constantes, de acuerdo a la Ecuación (3.4) de Mendez et al. [6]. En este contexto, G corresponde al factor de conversión análogo-digital, correspondiente a 2 [e^-/ADU]

4.2. Resultados experimentales

Ya habiendo mostrado las ganancias con respecto al error cuadrático en término de los límites fundamentales de los distintos enfoques de estimación, es importante constatar cuan alcanzables son estos resultados por los estimadores Bayesianos. Se ha dicho que el estimador Esperanza Condicional es el estimador que minimiza el error cuadrático medio, por lo que no hay otro estimador con mejor desempeño al analizar esta figura de mérito. No hay garantías teóricas de que se alcance la cota Bayesiana de Cramér-Rao, ya que no se cumple la Ecuación (2.14). De esta manera, solamente un análisis empírico puede explicar el comportamiento de este estimador en función del MSE.

4.2.1. Caso p.d.f a priori Uniforme

Al usar una p.d.f a priori uniforme, se obtiene el caso más parecido a lo que ocurre en caso paramétrico, donde todas las posiciones del objeto en el CCD son igualmente probables, lo que podría asociarse al esenario de no tener información de la posición de la estrella. Ya con el análisis hecho en la Ecuación (2.19), se sabe que esto es cierto, ya que la desigualdad de Jensen asegura que el mejor caso es que sean exactamente iguales, lo que sigue la idea intuitiva creada previamente.

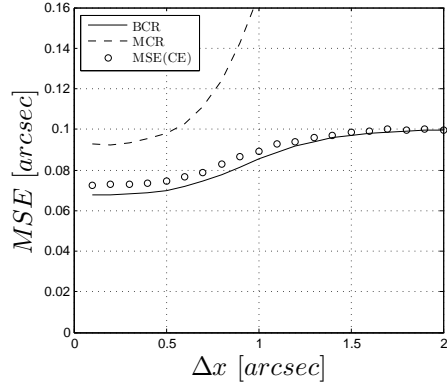
F(ADU)	EC (σ_{pix})	ML (σ_{pix})	LS (σ_{pix})	C-R (σ_{pix})	V-T (σ_{pix})
268	0.4308	0.423	0.423	0.406	0.4104
540	0.2136	0.212	0.212	0.209	0.2115
1612	0.0803	0.080	0.081	0.079	0.0799
3222	0.0458	0.0429	0.0437	0.045	0.0456
10002	0.0203	0.0212	0.0227	0.020	0.0202

Tabla 4.1: Valores de MSE de distintos estimadores y cotas, para distintos valores de flujo. En este caso, tanto la Esperanza Condicional como Cramér-Rao Bayesiano usan una p.d.f a priori uniforme.

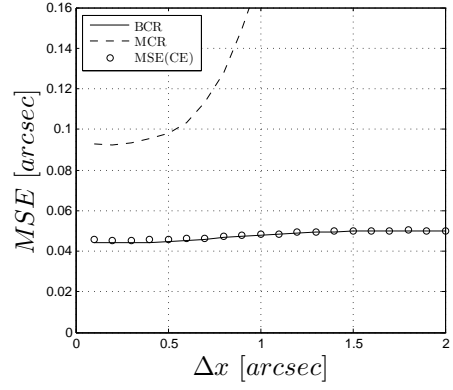
Al constatar los resultados obtenidos de BCR y el error cuadrático medio de la Esperanza condicional, se puede ver que en que, para un $\Delta x = 0.2$ [arcsec], *Background* fijo de 300 [ADU], los resultados son similares, variando de manera muy pequeña en los casos en que la relación señal-ruido es baja.

4.2.2. Caso p.d.f Normal

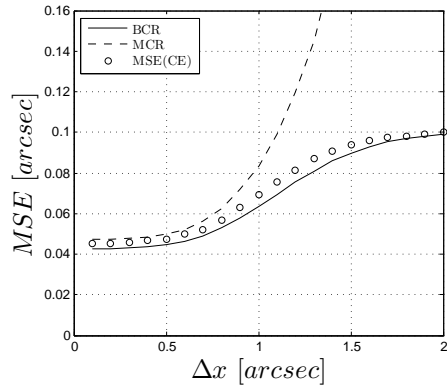
Al analizar los casos de $\sigma_{priori} = 0.1$ [arcsec] y $\sigma_{priori} = 0.05$ [arcsec], con un $FWHM = 1''$, se pueden ver mejoras significativas en las curvas de error cuadrático medio con respecto al tamaño del pixel. Es por lo mismo que ahora se toma como referencia este análisis para ver si efectivamente el caso de una p.d.f $N(\mu, \sigma_{priori})$ logra tener un comportamiento similar a BCR.



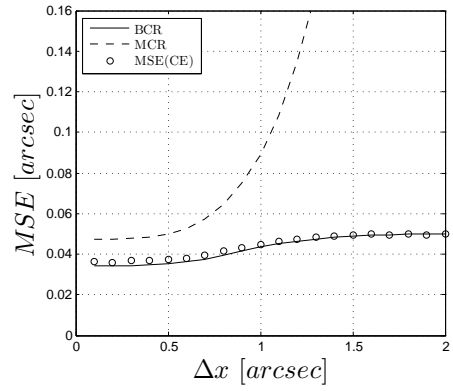
(a) SNR = 6 y $\sigma_{priori} = 0.1''$



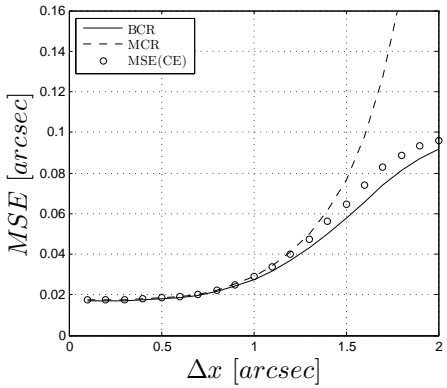
(b) SNR = 6 y $\sigma_{priori} = 0.05''$



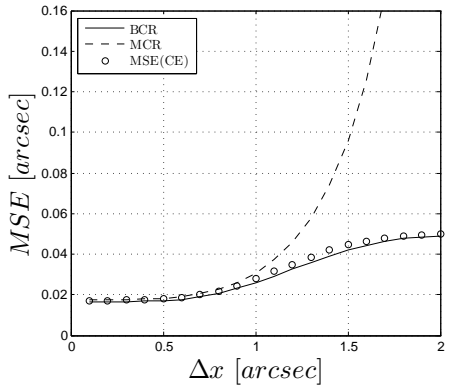
(c) SNR = 12 y $\sigma_{priori} = 0.1''$



(d) SNR = 12 y $\sigma_{priori} = 0.05''$



(e) SNR = 32 y $\sigma_{priori} = 0.1''$



(f) SNR = 32 y $\sigma_{priori} = 0.05''$

Figura 4.7: MSE v/s tamaño del pixel, para MSE de la EC y BCR con $FWHM = 1''$ y distintos valores de SNR y σ_{priori} .

En la Figura 4.2.2, se puede ver que efectivamente las curvas son similares, tendiendo el mismo comportamiento tanto en la región de interés ($\Delta x = 0.2 [arcsec]$) como en el caso de submuestreo de la estrella, donde el error cuadrático medio tiende al valor original de σ_{priori} . Tal como se había dicho previamente, las curvas no tienen garantías teóricas de ser iguales, lo cual se puede ver en este caso,

pero aún así, las diferencias entre ambas son pequeñas, dando una diferencia de 3% para el caso de un $\sigma_{priori} = 0.05 [arcsec]$ y un 7% en el caso $\sigma_{priori} = 0.1 [arcsec]$, ambos en el caso de un $SNR = 6$ lo que en terminos prácticos no representa una diferencia significativa, más aún en el caso que se están logrando ganancias de un 110% en el mejor de los casos estudiados. Con respecto a las otras relaciones de señal-ruido, se aprecian los mismos comportamiento, a pesar de que las ganancias van disminuyendo, tal cual lo que aprecia en el análisis de las cotas fundamentales.

Al realizar un ejercicio similar de la Sección 4.1.3, respecto a calcular el σ_{priori} necesario para que un flujo aparente obtenga el mismo MSE de la cota de Cramér-Rao que un flujo real utilizando la cota de Cramér-Rao Bayesiano, se puede constatar en las siguientes tablas el resultado. En la primera tabla se observa los resultados para un $FWHM = 1''$, en este caso se puede observar, más allá de lo visto en la Figura 4.6, cuan cercanos son los valores de Cramér-Rao Bayesiano con respecto al MSE de la Esperanza Condicional, teniendo nuevamente diferencias marginales con respecto al Error Cuadrático Medio.

$F_{par}(ADU)$	S/N	σ_{priori}	MSE(BCR $_{F=268, \sigma_{priori}}$)	MSE(C-E $_{F=268, \sigma_{priori}}$)
540	12	0.0547	0.0471	0.0486
1612	32	0.0176	0.0173	0.0174
3222	55	0.0098	0.0097	0.0098
10002	120	0.0051	0.0051	0.0051
30080	230	0.0032	0.0032	0.0032

Tabla 4.2: MSE en Análisis de Flujo aparente F_{par} , para la Cota de Cramér-Rao Bayesiana y el MSE de la Esperanza Condicional, con un $FWHM = 1''$.

Las tendencias no varán en el caso de un $FWHM = 2 [arcsec]$, aunque si se puede ver son necesarios valores de σ_{priori} menos ajustados como para obtener los mismos valores de flujo aparente.

F(ADU)	S/N	σ_{priori}	MSE(BCR $_{F=268, \sigma_{priori}}$)	MSE(C-E $_{F=268, \sigma_{priori}}$)
540	12	0.1458	0.1265	0.1310
1612	32	0.0442	0.0435	0.0435
3222	55	0.0227	0.0226	0.0226
10002	120	0.0086	0.0086	0.0086
30080	230	0.0046	0.0046	0.0046

Tabla 4.3: MSE en Análisis de Flujo aparente F_{par} , para la Cota de Cramér-Rao Bayesiana y el MSE de la Esperanza Condicional, con un $FWHM = 2''$.

Por lo tanto, al constatar las ganancias expuestas anteriormente en la Sección 4.3.1. con respecto solamente a los límites fundamentales se traducen en ganancias reales al utilizar el estimador Esperanza Condicional, para regímenes en donde se pueden tener ganancias. De manera que el estimador, además de ser óptimo, es cercano a la cota de Cramér-Rao Bayesiano y por tanto muchas de las ganancias de tener información a priori en un sistema se traducen en ganancias reales en el uso de estimadores basados en este enfoque.

Capítulo 5

Conclusiones y trabajos futuros

El área de la astronomía siempre ha tenido la constante necesidad de tener herramientas precisas para obtener un buen entendimiento del cosmos, esto ya sea a través de una buena tecnología como también buenos algoritmos y técnicas matemáticas para realizar sus fines. En esta Memoria de Título se buscó poner a disposición de esta disciplina una nueva aproximación al problema de la determinación de la posición del centroide en objetos luminosos, pero esta vez haciendolo con herramientas que dan garantías de ser óptimo desde el punto de vista del Error Cuadrático Medio.

El objetivo de esta memoria era precisamente ver cuantas ganancias se pueden obtener al usar el enfoque Bayesiano de estimación con respecto al enfoque tradicional paramétrico para el problema planteado, y se ha podido ver que en primera instancia, en lo que respecta a las cotas fundamentales de ambos enfoques, existe un espacio importante de ganancia en ciertos regímenes.

La primera ganancia que se puede observar en términos de la precisión es cuando se aumenta el tamaño del pixel. Debido a que el caso Bayesiano siempre estará acotado por el error a priori que se tenga del problema. Esto hace al método Bayesiano hace más robusto a condiciones donde las observaciones son menos informativas.

La ganancia más relevante se puede apreciar al estudiar un tamaño de pixel que normalmente se usa en aplicaciones astronómicas. Se puede apreciar que para valores de señal-ruido bajas y una información a priori ajustada, existen ganancias significativas con respecto al enfoque tradicional paramétrico, llegando incluso la Cota de Cramér-Rao Bayesiana a ser una quinta parte del valor de la esperanza de la Cota de Cramér-Rao. Esta tendencia se puede ver para distintos tamaños de estrellas, manteniendose esa tendencia.

En términos de las ganancias relativas, se hizo un estudio donde se analizó la ganancia de Cramér-Rao Bayesiano con respecto a la Esperanza de Cramér-Rao, para distintos valores de error a priori σ_{priori} . Del análisis de esas curvas se obtuvo que existe un compromiso entre obtener un buen resultado del error y sesgar el resultado imponiendo una precisión a priori muy alta, por lo

que no todos los valores de σ_{priori} son óptimos, sino que existe un conjunto acotado, el cual es importante tener en cuenta al diseñar el estimador final del problema astrométrico.

Un estudio hecho en esta memoria fue entender cuán certera debe ser la información previa del sistema, para que a través de un valor de flujo dado, se pueda tener un resultado (en términos del Error Cuadrático Medio) equivalente a un flujo al inicialmente mayor usando un caso de estimación paramétrica. El resultado de la Figura 4.6 da la posibilidad de estimar usando valores catálogos astronómicos como σ_{priori} para aumentar la precisión astrométrica.

Finalmente al apreciar el error cuadrático medio de la Esperanza Condicional, se puede ver que en los casos estudiados el comportamiento es similar al de la Cota de Cramér-Rao Bayesiana, no habiendo más allá de un 7% de diferencia entre ambas. Esto es significativo, ya que si bien no existían garantías teóricas de que fueran iguales, los resultados empíricos muestran notables semejanzas que servirán para poder hacer, a través de la Cota de Cramér-Rao Bayesiana, un diseño bastante preciso de lo que ocurrirá con el error del estimador.

Si bien el estudio se realizó abordando una amplia cantidad de factores, la implementación de la varianza del estimador no se hizo pensando en términos de eficiencia sino más bien prácticos, así que hay todo un espacio de mejora en lo que tiene que ver con la cantidad de iteraciones y el tiempo de convergencia del algoritmo, el cuál se puede mejorar con algoritmos que apuntan precisamente a solucionar esas falencias, como lo puede ser el algoritmo de *Importance sampling*, el cuál, al adjudicar ciertos pesos de acuerdo a la formas de las funciones de distribución de probabilidad, mejora considerablemente la eficiencia del proceso.

Finalmente, queda pendiente este estudio con datos experimentales no simulados, donde se pueda confrontar los resultados aquí expuestos con la realidad.

Bibliografía

- [1] B. Z. Bobrovsky, E. Mayer-Wolf, and M. Zakai. Some classes of global cramer-rao bounds. *The Annals of Statistics*, 15(4):pp. 1421–1438, 1987.
- [2] Justin Dauwels. Computing bayesian cramer-rao bounds. pages 425–429, 2005.
- [3] Richard D Gill and Boris Y Levit. Applications of the van trees inequality: a bayesian cramer-rao bound. *Bernoulli*, pages 59–79, 1995.
- [4] P Jakobsen, P Greenfield, and R Jedrzejewski. The cramer-rao lower bound and stellar photometry with aberrated hst images. *Astronomy and astrophysics*, 253(1):329–332, 1992.
- [5] Steven M Kay. *Fundamentals of statistical signal processing: estimation theory*. Prentice-Hall PTR, 2010.
- [6] Rene Mendez, Jorge Silva, and Rodrigo Lobos. Analysis and interpretation of the cramer-rao lower-bound in astrometry: One dimensional case. *arXiv preprint arXiv:1304.6361*, 2013.
- [7] Rene A Mendez, Jorge F Silva, Rodrigo Orsotica, and Rodrigo Lobos. Analysis of the cramer-rao lower uncertainty bound in the joint estimation of astrometry and photometry. *arXiv preprint arXiv:1407.3753*, 2014.
- [8] Kenneth J Mighell. Stellar photometry and astrometry with discrete point spread functions. *Monthly Notices of the Royal Astronomical Society*, 361(3):861–878, 2005.
- [9] Todd K Moon and Wynn C Stirling. *Mathematical methods and algorithms for signal processing*, volume 1. Prentice hall Upper Saddle River, 2000.
- [10] Harry L Van Trees. *Detection, estimation, and modulation theory*. Wiley. com, 2004.
- [11] E Weinstein and AJ Weiss. Lower bounds on the mean square estimation error. *Proceedings of the IEEE*, 73(9):1433–1434, 1985.

- [12] Ehud Weinstein and Anthony J Weiss. A general class of lower bounds in parameter estimation. *Information Theory, IEEE Transactions on*, 34(2):338–342, 1988.
- [13] Kim A Winick. Cramer-rao lower bounds on the performance of charge-coupled-device optical position estimators. *JOSA A*, 3(11):1809–1815, 1986.
- [14] TS Zaccheo, RA Gonsalves, SM Ebstein, and P Nisenson. Estimating the cramer-rao bound for restored astronomical observations. *The Astrophysical Journal*, 439:L43–L45, 1995.