Practica I - Parte 1

Estudio Experimental de Distribuciones de Probabilidad

Introducción a las Ciencias de la Tierra y el Espacio I - 2011

Índice

1. Introducción

En la física, como en cualquier ciencia experimental, los fenómenos que se analizan deben poder observarse y medirse. Estas medidas se presentan bajo la forma de números que expresan el valor de determinadas magnitudes y lo que esperamos de ellos es que nos permitan confirmar o no un modelo, desarrollar un trabajo tecnológico, etc. Al hacer una medida verificamos que no se obtienen los resultados deseados en forma tan directa. Algunas veces la medida está sujeta a errores que, frecuentemente, no podemos eliminar pues son propios del proceso de medición. En otros casos, la magnitud medida tiene un carácter esencialmente estadístico por lo que no podemos hablar de valor exacto de la magnitud. Cuando queremos medir una magnitud repetimos algunas veces la observación para tener una idea del error que podemos estar cometiendo en cada medición. Sin embargo, la diferencia entre los varios datos obtenidos nos dará una idea del error experimental si cada observación fuera independiente de todas las otras. Que significa esto? Para que en una medida los datos sean independientes del dispositivo experimental es necesario que la obtención de determinado valor no interfiera en la obtención de cualquier otro. Para que haya una independencia completa entre las diversas observaciones, sería (en teoría) necesario que cada dato sea obtenido con un equipo diferente, por otro observador, etc. Esto es impracticable y carente de sentido desde el punto de vista experimental. Sin embargo el tema de la independencia de los datos del proceso de medición es un aspecto que el investigador debe considerar seriamente. En este curso estudiaremos algunos métodos para realizar un tratamiento de datos experimentales que nos permita garantizar la calidad de la medida obtenida.

2. Proceso de Medición

Comenzaremos por analizar el proceso de medición; en él intervienen necesariamente tres objetos: el sistema objeto, el cual queremos medir; el sistema de medición, o sea el aparato con el que mediremos y por último el sistema de comparación que definiremos como unidad. Por ejemplo, en el proceso llamado "medición de longitud" intervienen

1. El objeto cuya longitud se desea medir
2. El instrumento (por ejemplo una regla)
3. La unidad (cierta escala marcada en la regla o en cierta barra patrón)

Cada proceso de medición define lo que se llama magnitud física. Si dos procesos definen la misma magnitud física, son equivalentes. El resultado de un proceso de medición es un número real, que se denomina valor de la magnitud. Para una magnitud dada, su valor debe ser independiente del proceso particular de medición, dependiendo únicamente de la unidad elegida. Haremos a continuación un breve repaso dal tipo de errores que usualmente se cometen cuando se realiza una experiencia.

2.1. Errores Sistemáticos

Son aquellos que afectan a todos los datos por igual. Pueden ser debido a una mala calibración del instrumento de medida o un error de lectura del operador. Estos errores se llaman sistemáticos justamente porque pueden ser eliminados mediante un correcto diseño del montaje experimental.

2.2. Errores Estadísticos

Son debido a los diferentes fenómenos de naturaleza aleatoria que ocurren dentro del experimento. Para una discusión más profunda acerca de errores de naturaleza aleatoria ver el apéndice acerca de *Distribuciones de Probabilidad y Distribución Gaussiana*.

2.3. Error Absoluto

Para la medida de la magnitud física X, el error absoluto ΔX es el error asignado como resultado del proceso de medición. El resultado de la medida se expresara de la siguiente forma

$$X \pm \Delta X \tag{2.1}$$

Lo que significa que el *valor verdadero* X de la magnitud medida pertenece al siguiente intervalo

$$X \in [X + \Delta X, X - \Delta X]$$

2.4. Apreciación

Es una característica del instrumento de medida y es el intervalo más pequeño que puede registrar el instrumento. En los instrumentos de medida con escalas graduadas, la apreciación es la medida que hay entre dos marcas consecutivas de la escala. Por ejemplo, si medimos una temperatura con un termómetro de mercurio graduado en $1/10$ de grado, la apreciación del mismo será de $1/10°$.

2.5. Estimación

Es una acción que tomamos frente a una medida que realizamos con un instrumento cuya escala está toda a la vista. Radica en poder subdividir aún más dicha escala (a un nivel más fino que la propia apreciación del instrumento), y así ganar precisión en la medida final. Un ejemplo de esto es lo que ocurre cuando un fiel cae en la mitad de dos intervalos consecutivos de la escala. Podemos subjetivamente dividir ese intervalo a la mitad y entonces afinar nuestra medida. La estimación siempre la realizamos subdividiendo la apreciación.

2.6. Error Relativo

Nos da una idea de la calidad de la medida. El mismo se denota con la letra ε y se define como el cociente entre el error absoluto de la medida y la medida misma. Si el error absoluto es menor que la medida, el error relativo es menor que la unidad. Evidentemente en el proceso de medición se busca optimizar el resultado experimental y por lo tanto disminuir el error relativo.

$$\varepsilon_x = \frac{\Delta X}{X} \tag{2.2}$$

2.7. Error Relativo Porcentual

Es el mismo error relativo anterior pero expresado en porcentaje o escala de 0 a 100, en vez de 0 a 1 como el error relativo. Se define de la siguiente manera

$$\varepsilon_{x\,\%} = \varepsilon_x \cdot 100 \tag{2.3}$$

Podemos decir en general que si obtenemos un error menor al 10 % el modelo utilizado ajusta razonablemente bien los resultados experimentales. En caso contrario será necesario realizar correcciones en el modelo. Por ejemplo, si estamos estudiando el movimiento de un cuerpo sometido a fuerzas de rozamiento y consideramos un modelo que supone velocidad constante, seguramente obtendremos un porcentaje de error muy grande. En este caso hay al menos dos alternativas: o adaptamos la experiencia para eliminar el rozamiento, por ejemplo trabajamos en una mesa de aire comprimido, o adaptamos el modelo incluyendo las variaciones en la velocidad. En la práctica, en general, hay que optar por soluciones que contemplen ambas modificaciones.

2.8. ¿Cómo asignar el error a una magnitud medida?

En este punto, los criterios a seguir suelen ser muy diversos. Si se consultan cinco libros diferentes sobre errores de mediciones, probablemente se encuentren cinco criterios distintos para asignar un error a la lectura. En clase se adoptará el siguiente criterio.

- Instrumentos Analógicos: Se adoptará la estimación de la lectura como el error Δx de la medida x. Por ejemplo, si se mide con un termómetro que aprecia al grado una temperatura de 15 grados y medio. Si el usuario tiene mucha práctica quizá pueda estimar 1/10 de grado, entonces anotará: $T = (15{,}5 \pm 0{,}1)^{\circ}C$. Si no se siente con tanta confianza, deberá estimar un poco menos (1/5 o 1/2 de grado).

- Instrumentos Digitales: Los buenos *testers* incluyen en su manual un instructivo sobre cómo calcular el error en la lectura. Generalmente es un porcentaje de ésta, más un valor fijo. No obstante, el error calculado de esta forma afecta sólo el último dígito de la lectura (si afectaran los dos últimos dígitos habría que dudar en usar ese tester). Los *testers* de mediana calidad y los demás instrumentos digitales de uso masivo (cronómetros, computadora, etc) no traen explícitamente la forma de calcular el error. En vista de ello se adoptará para todos los instrumentos digitales, la apreciación del instrumento digital como el error. Esto significa que si se lee un valor de voltaje con un tester digital y se obtiene una lectura de 4,58v, deberá enunciarse la medida con su error de esta forma: $V = (4{,}58 \pm 0{,}01)v$

3. Conceptos Básicos de Probabilidad

Para poder trabajar con variables de tipo aleatorias y analizar los resultados experimentales obtenidos en el proceso de medición tendremos que introducir una descripción estadística del sistema físico. Es decir será necesario introducir el concepto de *probabilidad*.

3.1. Variable Aleatoria

Supongamos que medimos cierta magnitud física x con un instrumento extremadamente sensible, de forma que la apreciación de la lectura es despreciable en comparación al valor de la medida. En estas condiciones, es un hecho experimental que si medimos muchas veces la magnitud x en iguales condiciones del sistema, no obtendremos siempre el mismo valor x. Decimos entonces que la magnitud x es una *variable aleatoria*.

- La probabilidad que ocurra un evento r en un total de eventos N, esta definida como el cociente entre el número de eventos favorables N_r sobre los eventos totales N.

$$P_r = \frac{N_r}{N} \qquad \text{con} \quad N \to \infty \tag{3.1}$$

- La probabilidad que ocurran el evento r o el evento s, esta dada por la suma de las probabilidades de r y s.

$$P_{r \circ s} = P_r + P_s \tag{3.2}$$

- La probabilidad que ocurran el evento r y el evento s, al mismo tiempo, esta dada por el producto de sus probabilidades. Para que esto se cumpla, r y s deben ser eventos *estadísticamente independientes*.

$$P_{rs} = P_r \times P_s \tag{3.3}$$

- La probabilidad p cumple que $p \in [0,1]$. Donde la probabilidad es 0 corresponde al evento que nunca ocurrirá y la probabilidad 1 al que sí o sí ocurrirá. Si tenemos muchos eventos posibles, la probabilidad que ocurra alguno de ellos, corresponde con la suma de las probabilidades (visto anteriormente), por lo tanto la suma de las probabilidades de todos los eventos posibles es

$$\sum_{r=1}^{N} P_r = 1 \tag{3.4}$$

4. Histogramas

Una forma de representar una colección de datos obtenidos en una medida experimental es la construcción de una gráfica denominada *histograma*. Consideremos la siguiente experiencia: una bolita que se mueve por una rampa. Nos interesa determinar el tiempo que demora en recorrer una cierta distancia. Para ello se tira la bolita varias veces y se determina el tiempo en cada caso. Observe que el histograma muestra gráficamente el número de veces que se obtuvo una medida entre un valor y otro (también llamado casillero o *bin*), o sea, es un gráfico de la acumulación de valores. En la figura (4.1), se muestran cuatro histogramas construidos a partir de datos obtenidos experimentalmente. En las figuras (a), (b) y (c) cada serie tiene 20 muestras, en tanto que en la figura (d), se adquirieron 200 muestras. Hay varios puntos interesantes a resaltar.

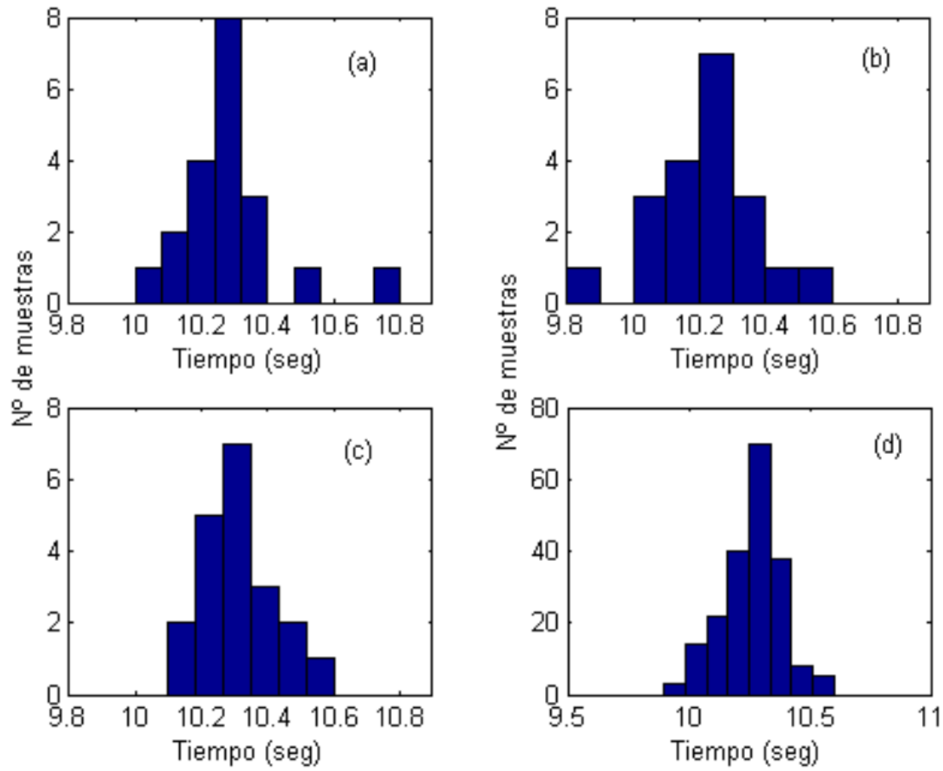

Figura 4.1: Para la serie de 200 datos, representada en la figura (d) se observa que la distribución tiende a ser simétrica

- No hay diferencias esenciales en la forma en que se distribuyen los datos en cada una de las series.
- En los todos los casos los valores de la serie se distribuyen alrededor de un mismo valor $t_o = 10{,}3s$. Esta distribución de datos es básicamente la misma en cada serie. Sin embargo se observa que, las tres primeras distribuciones no presentan simetría en torno a t_o, ver figura (4.1).
- La forma del histograma puede variarconsiderablemente si consideramos bins distintos. Esto es si ampliamos o reducimos el tamaño de los casilleros en el que contamos la cantidad de veces que cae una medida. Este ejemplo se ilustra en la figura (). Generalmente basta con tomar unos pocos bins (10) para observar tendencias de acumulación generales como ser comportamientos unimodales, bimodales o multimodales. Si disponemos de una gran cantidad de medidas de una variable aleatoria es posible afinar el tamaño de los bins y observar la curva característica gaussiana de la distribución de los valores.

Intuitivamente podemos decir que rehaciendo las series de medidas experimentales, bajo las mismas condiciones, e independientemente del número de datos adquiridos, obtendremos histogramas similares, es decir similares funciones de distribución. Lo interesante de esta última observación es que nos va a permitir asociar a cada serie de datos experimentales una función de distribución que va a representar la probabilidad de que un determinado dato de la serie pertenezca a cierto intervalo de medidas.

4.1. Histogramas en Matlab

La función *hist*, permite realizar histogramas de una forma muy fácil. Para esto debemos definir un vector V cuyos elementos sean los N valores de una medida X cualquiera.

```
Y=[1 1.1 0.9 1 1 1 0.8 1.2 1.2 1.1 1 1 1 1 0.7 0.8 0.9 0.9 0.9 1 1.1 1.1 1 1 1]
```

El histograma correspondiente se obtiene ejecutando la siguiente tarea. Observe que por defecto Matlab utiliza 10 *bins* o casilleros entre el primer valor y el último.

```
hist(Y)
```

Si se quiere generar un histograma con 4 bins

```
hist(Y,4)
```

Si solamente se quiere el vector de la cantidad de veces que se obtuvieron los valores en los respectivos bins

```
N=hist(Y,4)
```

En la figura (4.2) se muestran diferentes histogramas para la misma serie de valores de la magnitud X. Observe que el aspecto de los histogramas varía considerando distintos tamaños de bins. Observe que para los histogramas de 8 y 10 bins, comienzan a generarse vacíos. Un vacío significa que entre un cierto valor y otro, no se obtuvo ninguna medida. Cuando los vacíos en un histograma aparecen uniformemente a lo largo del mismo (caso de 10 bins) significa que estamos considerando bins demasiado pequeños. En el caso contrario observe que de tomar muy pocos bins, el histograma tiende a una sola barra donde ahí estarán todas las medidas de la magnitud X. El histograma de 6 bins resulta el más adecuado, ver figura (4.2)

4.2. Ejercicio (obligatorio)

4.2.1. Diámetro del cráter lunar *Tycho*

En base a la figura que se adjunta, realizar dos series de medidas para la distancia del pico central al borde del cráter de manera de recorrer toda la circunferencia en sentido horario u anti-horario. El diámetro del cráter es de $85km$.

4.2.2. Cálculo en Matlab

Iniciar Matlab y setear el directorio de trabajo en la carpeta de trabajo (la misma carpeta donde se encuentra la imagen del cráter). Cargar la imagen a la variable A.

```
A=imread('tycho.jpg','jpg')
```

Luego abrir la imagen como una figura en Matlab con el siguiente comando

```
imagesc(A)
```

En este momento se desplegará la imagen en una ventana de una figura. Observe que la relación de aspecto entre los ejes no es la misma, esto ocasiona que la imagen se vea ensanchada o alargada. Para evitar esto y que la imagen se visualice correctamente hay que ejecutar el siguiente comando

```
axis equal
```

Ahora la imagen esta correctamente visualizada y ya podemos comenzar con las medidas del centro del cráter y los radios. Para realizar esto definimos las coordenadas del centro con la función *ginput.*

```
[xc,yc]=ginput
```

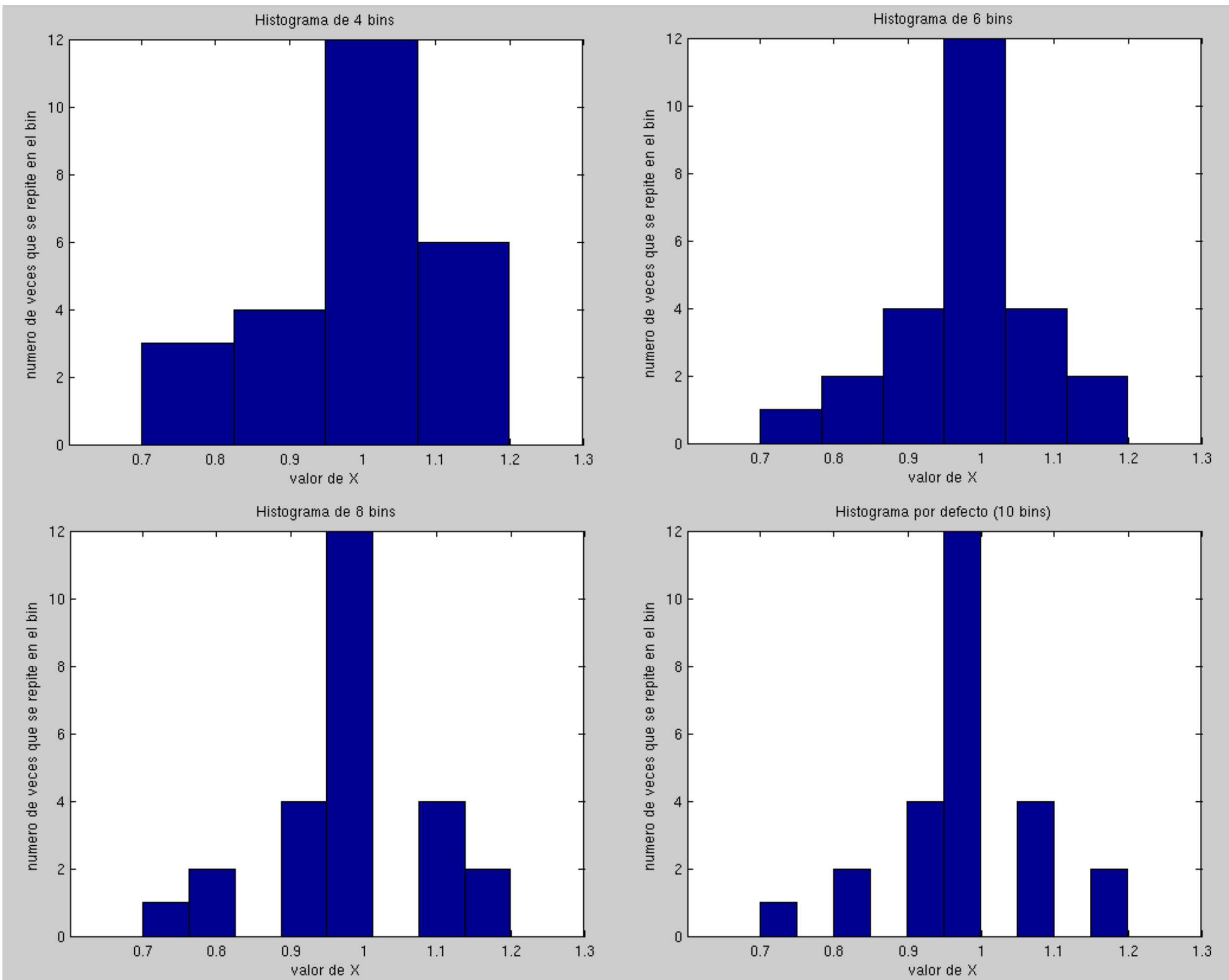

Figura 4.2: SUPERIOR-IZQUIERDA: Histograma de 4 bins. SUPERIOR-DERECHA: Histograma de 6 bins. INFERIOR-IZQUIERDA: Histograma de 8 bins. INFERIOR-DERECHA: Histograma de 10 bins. Matlab utiliza 10 bins por defecto si el usuario no indica el número.

En este instante el cursor del mouse sobre la imagen se transformará en una cruz. Para registrar las coordenadas del centro del cráter (pico central), hacemos un click en el lugar deseado. Luego para finalizar la tarea apretar la tecla *enter*. De manera similar, registraremos las coordenadas de varios puntos (20 y 50) en 2 vectores x e y.

```
[x,y]=ginput
```

Luego de hacer las medidas correspondientes y apretar la tecla enter para finalizar, tendremos los valores de x_c, y_c y los vectores x e y con las coordenadas de los puntos del borde del cráter. Para calcular las distancias del pico central al borde aplicamos *Pitágoras* con la siguiente operación

```
D=sqrt((x-xc).^2+(y-yc).^2)
```

Luego el vector D contiene todas las distancias del borde del cráter al pico central. En la primera serie se tomarán 20 medidas a guardar en un archivo llamado *datos20.txt* y en la en la segunda serie se tomarán 50 medidas a guardar en un archivo llamado *datos50.txt*. Ambos archivos se cargaran en Matlab y se harán los histogramas correspondientes, ver subsección (4). Se discutirá la forma de los mismos y se obtendrá un valor para el valor promedio de la medida en cada caso, asi como el error asociado.

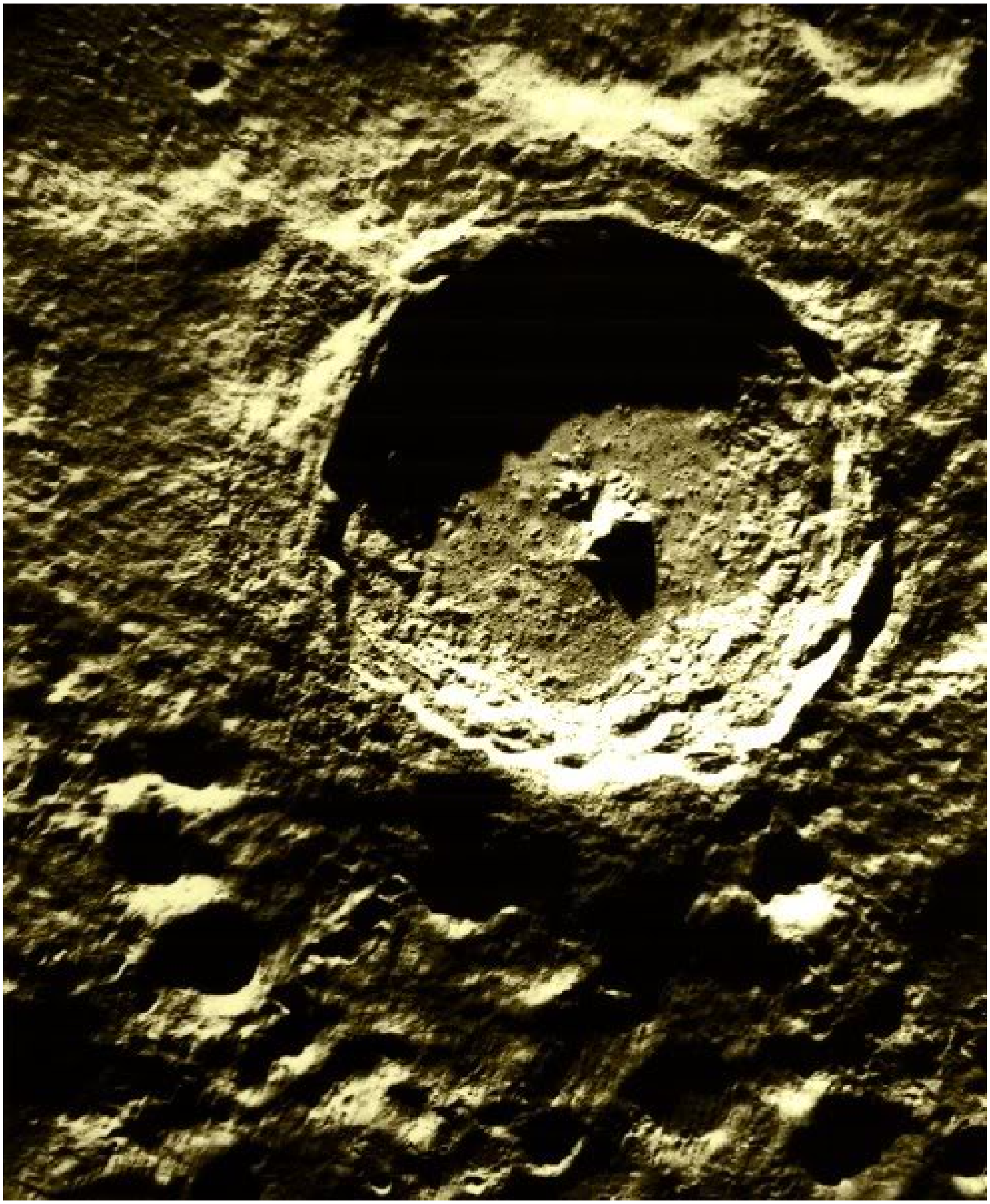

Figura 4.3: Cráter lunar de *Tycho*. Su diámetro es de $85km$.