Capítulo 4

Recopilación de Datos y Prueba de Bondad de Ajuste

Capítulo 4

Recopilación de Datos y Prueba de Bondad de Ajuste

Los objetivos de este capítulo son los siguientes:

- Conocer los conceptos básicos sobre prueba de hipótesis.
- Presentar algunas pruebas de bondad de ajuste para probar que tipo de comportamiento tiene las distribuciones.
- Comprender técnicas de Recopilación de datos.

4.1 Introducción

Una vez definida nuestra variable de estudio y recopilados los datos para que los datos sean útiles, necesitamos organizar nuestras observaciones de modo que podamos distinguir patrones y llegar a conclusiones lógicas. Los especialistas en estadísticas seleccionan sus observaciones de manera que todos los grupos relevantes estén representados en los datos.

Mientras el sistema se desarrolla de acuerdo con un proceso aleatorio, debemos colectar datos y surgirá el problema de comprender el comportamiento de dicho sistema para poder luego simularlo en el computador. No son suficientes las estimaciones de esperanza matemática y desviación estándar para este propósito.

Para ello es necesario a partir de una hipótesis investigar qué tipo de distribución probabilística corresponde al comportamiento de nuestro experimento. De esta manera, podremos generar números aleatorios que correspondan a dicha distribución y comportamiento.

En este capítulo presentamos algunas reglas prácticas que ayudan al analista a lograr este fin, y dos métodos de pruebas de hipótesis comúnmente empleados; la prueba de Bondad de Ajuste Chi-Cuadrada y la Prueba no paramétrica Kolmogorov-Smirnov.

4.2 Recopilación de Datos

Una vez seleccionadas las variables involucradas en nuestra investigación y la muestra adecuada de acuerdo con nuestro problema en estudio y las hipótesis formuladas, la siguiente etapa consiste en recolectar los datos pertinentes sobre las variables involucradas en la investigación.

Recolectar datos implica tres actividades estrechamente relacionadas entre sí:

- **Seleccionar un instrumento de medición,** de los disponibles en el estudio o desarrollar uno para la recopilación. Este instrumento debe ser válido y confiable, de lo contrario no podremos tomarlos como base para nuestras conclusiones.

- **Aplicar ese instrumento de medición,** es decir, obtener las mediciones y observaciones de las variables que son de interés para nuestro estudio.

- **Preparar las mediciones obtenidas,** para que puedan analizarse correctamente, esta actividad se llama codificación.

4.2.1 ¿Qué significa “Medir”?

Una definición bien sencilla sería: [1] “Asignar números a objetos y eventos de acuerdo con reglas”, esta definición es apropiada para las ciencias físicas, no así para ciencias sociales, ya que varios fenómenos que son medidos en estas no pueden caracterizarse como objetos o eventos, puesto que son demasiados abstractos para ello. La disonancia cognoscitiva, la alienación, el producto nacional interno y la credibilidad son conceptos demasiado abstractos para ser considerados “cosas que pueden verse o tocarse” (definición de objeto) o solamente como “resultados, consecuencias o producto” (definición de evento).

Este razonamiento nos hace definir medición como **el proceso de vincular conceptos abstractos con indicadores empíricos**, este proceso que se realiza mediante un plan explícito y organizado, para clasificar, (y frecuentemente cuantificar) los datos disponibles, **los indicadores** en término del concepto que el investigador tiene en mente. En este proceso el instrumento de medición o recopilación de datos juega un papel muy importante, sin él no hay observaciones clasificadas.

La definición sugerida incluye dos consideraciones: La primera es desde el punto de vista empírico y donde el centro de atención es la respuesta observable, sea una alternativa de respuesta marcada en un cuestionario, una conducta grabada, vía de observación o una respuesta dada al entrevistador.

La segunda es desde una perspectiva teórica y se refiere a que el interés se sitúa en el concepto subyacente no observable que es representado por la respuesta.

Un instrumento de medición adecuado es aquel que registra datos observables que representan verdaderamente los conceptos o variables que el investigador tiene en mente.

[1] Metodología de la Investigación. Pág. 234

4.2.2 Requisitos de un Instrumento de Medición

Toda medición o instrumento de recopilación de datos debe reunir dos requisitos esenciales: **confiabilidad** y **validez**.

4.2.2.1 Confiabilidad

La confiabilidad en un instrumento de medición se refiere al grado en que su aplicación repetitiva al mismo sujeto u objeto produce iguales resultados. Por ejemplo, si se midiera en este momento la temperatura ambiental mediante un termómetro que indica 22 ^{o}C. Un minuto más tarde se consulta al mismo termómetro e indica 5 ^{o} C. Tres minutos después marca 40^{0} C. Este no es un instrumento confiable, su aplicación repetida produce resultados diferentes.

4.2.2.2 Validez

La validez en términos generales se refiere al grado en que un instrumento realmente mide la variable que pretende medir. Por ejemplo, un instrumento válido para medir la inteligencia debe medir la inteligencia y no la memoria. Una prueba sobre conocimiento de historia debe medir eso y no-conocimiento de literatura histórica. La situación no es tan simple cuando tratamos las variables como la motivación, la calidad del servicio a los clientes, la actitud hacia un candidato político y menos aún con sentimientos y emociones, así como con diversas variables con las que se trabaja en las ciencias sociales.

Evidencias que están relacionadas con la Validez:

- Evidencia relacionada con el contenido.
- Evidencia relacionada con el criterio.
- Evidencia relacionada con el constructo[2].

Validez Total = Validez de Contenido + Validez de Criterio + Validez de Constructo

- **Validez de Contenido**

Se refiere al grado en que un instrumento refleja un dominio específico de contenido de lo que se mide. Por ejemplo, una prueba de operaciones aritméticas no tendrá validez de contenido si solamente se considera la resta.

- **Validez de Criterio**

La validez de criterio establece la validez de un instrumento de medición comparándola con algún criterio externo, este criterio es un estándar con el que se juzga la validez del instrumento; entre más se relacione los resultados del instrumento de medición con el criterio, la validez del criterio será mayor. Por ejemplo, un investigador valida un examen sobre manejo de aviones, mostrando la exactitud con que el examen predice que tan bien un grupo de pilotos puede operar un aeroplano.

---

[2] Un constructo es una variable medida y que tiene lugar dentro de una teoría o esquema teórico.

- **Validez de Constructo**

Es probablemente la más importante sobre todo desde una perspectiva científica y se refiere al grado en que una medición se relaciona consistentemente con otras mediciones de acuerdo con hipótesis derivadas teóricamente y que conciernen a los conceptos (o constructos) que están siendo medidos.

- **La validez de constructo incluye tres etapas:**

1. Se establece y especifica la relación teórica entre los conceptos (sobre la base del marco teórico).
2. Se correlacionan ambos conceptos y se analiza cuidadosamente la correlación.
3. Se interpreta la evidencia empírica de acuerdo con el nivel en que clarifica la validez de constructo de una medición en particular.

El proceso de validación de un constructo está vinculado con la teoría. No es posible llevar a cabo la validación de constructo, a menos que exista un marco teórico que soporte a la variable en relación con otras variables.

Así la validez de un instrumento de medición se evalúa sobre la base de tres tipos de evidencia. Entre mayor evidencia de validez de contenido, validez de criterio y validez de constructo tenga un instrumento de medición, este se acerca más a representar las variables que pretende medir.

Cabe agregar que un instrumento de medición puede ser confiable pero no necesariamente válido (Un aparato, por ejemplo, puede ser consistente en los resultados que produce, pero no medir lo que pretende).

Factores que afectan la confiabilidad y la validez:

- La improvisación.
- Datos perdidos.
- Mala selección de la muestra.

4.3 Técnicas de Recopilación de Datos

Un investigador se vale de técnicas o métodos para la recopilación de datos, algunas formas son:

- Cuestionarios.
- Observación.
- Pruebas.
- Inventarios estandarizados.
- Sesiones con profundidad.

- Archivos.
- Otros métodos de recopilación.

4.3.1 Proceso para construir un Cuestionario

El cuestionario es uno de los métodos de recopilación de datos más usados, los mismos tienen pasos a seguir para que sean eficaces.

- Revisión de la literatura de cuestionarios que midan las mismas variables que pretenderemos medir en la investigación.
- Evaluar la validez y la confiabilidad de cuestionarios anteriores.
- Consultar con expertos o personas familiares al tema investigado.
- Entrenar encuestadores en caso de que se requieran.
- Aplicar.

4.3.2 Pasos para construir un Sistema de Observación

- Definir con precisión el universo de aspectos, eventos o conductas a observar.
- Extraer una muestra representativa.
- Establecer y definir las unidades de observación.
- Establecer y definir las categorías y subcategorías de observación.

Definir con precisión el universo de aspectos, eventos o conductas a observar.

Por ejemplo, si nuestro interés es observar los recursos con que cuentan las escuelas de un distrito escolar debemos definir lo que concebimos como “recurso escolar”. Un universo podría ser el comportamiento verbal y no verbal de un grupo de alumnos durante un semestre. Otro universo sería las conductas de un grupo de trabajadores durante sus sesiones en círculos de calidad o equipos para la calidad, en un periodo de un año.

Extraer una muestra representativa de los aspectos, eventos o conductas a observar.

Un repertorio suficiente de conductas para observar.

Establecer y definir las unidades de observación.

Por ejemplo, cada vez que se presenta una conducta agresiva, cada minuto se analizará si el alumno está o no atento a clase, durante dos horas al día (7:00 a 9:00 horas), los números de personas que leyeron el tablero de avisos de la compañía, etcétera.

Establecer y definir las categorías de observación.

Es importante determinar con claridad las categorías, en caso de ser necesario, ya que existen muchas pruebas para datos categorizados.

4.3.3 Pruebas e Inventarios Estandarizados

En la actualidad existe una amplia diversidad de pruebas e inventarios desarrollados por diversos investigadores para medir un gran número de variables. Estas pruebas tienen su propio procedimiento de aplicación, codificación e interpretación y se encuentran disponibles en diversas fuentes secundarias y terciarias, así como en centros de investigación y difusión de conocimiento. Hay pruebas para medir habilidades y aptitudes, la personalidad, los intereses, la motivación, el desempeño, los valores y el clima laboral de una organización.

Cuando se utilicen como instrumento de medición es conveniente que se seleccione una prueba desarrollada o adaptada por algún investigador para el mismo contexto de nuestro estudio y que sea válida y confiable.

4.3.4 Sesiones en Profundidad

Un método de recopilación de datos cuya popularidad ha crecido son las sesiones en profundidad. Se reúne un grupo de personas y se trabaja con este en relación con las variables de la investigación. Pueden realizarse una o varias reuniones. El procedimiento usual es el siguiente:

- Se seleccionan los tipos de personas que habrán de participar en las sesiones.
- Se detectan los tipos de personas elegidas.
- Se invita a las personas a la sesión.
- Se lleva a cabo cada sesión. Cada sesión debe efectuarse en un lugar cómodo, confortable y silencioso, ya que el sujeto debe sentirse relajado.
- Se lleva a cabo cada sesión, el conductor debe ser una persona entrenada en el manejo de grupos y crear un ambiente o clima de confianza. Durante esta sesión pueden pedirse opiniones, hacer preguntas, administrar cuestionarios, discutir casos, intercambiar puntos de vista y valorar diversos aspectos.
- Elaborar el reporte de sesión, el cual incluye datos sobre el participante, fecha y duración de la sesión, información completa del desarrollo de la sesión. Actitud y comportamiento de los participantes hacia el conductor y la sesión en sí.
- Análisis correspondientes.

4.3.5 Otras Formas de Recopilación de Datos

- Archivos que contengan los datos.
- Análisis secundarios o datos recaudados por otros investigadores
- Análisis de redes.
- Combinación de dos o más instrumentos de recopilación de datos.

4.4 Síntesis de los Datos

Los datos en bruto y las largas listas de números, no son de mucha utilidad. Se les debe sintetizar o resumir de manera que se les pueda entender y utilizar. Imagínese que se encuentra frente al balance de cuenta corriente de un banco, impreso por una computadora, en él aparecen los límites de crédito y los pagos recientes de cada uno de los usuarios de la tarjeta VISA; nadie podría encontrarle sentido a esa pila de datos. La síntesis es el primer paso en el análisis de los datos.

Luego de reunidos los datos, se deben representar gráficamente. Algunos diagramas y gráficas pueden hacer simple algunos aspectos importantes entre los datos. Por ejemplo, un diagrama que muestre los porcentajes de los poseedores de las tarjetas VISA, que han pagado alrededor del 10% de sus deudas pendientes; y otro porcentaje de los que han pagado el 20% y así sucesivamente hasta llegar al 100%.

Un diagrama que indica la frecuencia con que las diferentes partes de una tostadora de pan, fue causa de una reparación en garantía señalaría muy claramente los problemas de calidad. El proceso de dar sentido a los datos normalmente comienza con una gráfica.

El último paso al reunir los datos consiste en determinar cualquiera forma o patrón que están presentes.

Queremos buscar entre ellos la asimetría y los valores atípicos, valores poco usuales rarezas que difieren considerablemente en los datos. Por ejemplo, un diagrama que represente la riqueza monetaria de las familias no será simétrico; la mayor parte de las familias tendrán un valor muy pequeño, otras tendrán valores moderadamente grandes, y otros valores verdaderamente grandes. Si no tomamos esto en cuenta y solo consideramos el promedio de las riquezas familiares, nuestra apreciación seria errónea.

Es importante conocer algunas técnicas para sumarizar o describir rasgos importantes del conjunto inicial de datos. Sin embargo, es importante mencionar que las técnicas para sumarizar datos implican la pérdida de cierta información contenida en ellos.

Datos Agrupados

Un método para lograr que los datos sean más manejables es agruparlos de acuerdo con clases. Se sumarían tabulando los números que caen dentro de cada clase.

A este tipo de tabla formada se le llama tabla de distribución de frecuencia y por lo general, tiene una buena y completa imagen de los datos.

Un ejemplo de tabla de distribución para datos agrupados relacionados al tiempo de espera de clientes es:

Tiempo de Espera (segundos)	Números de Clientes
0-20	21
20-40	35
40-60	42
60-80	35
80-100	19
100-120	10
>120	10

El número en la columna derecha denota la cantidad de clientes que caen dentro de cada clase. Los valores de la izquierda definen el rango de valores en cada clase y se refiere a los límites de éstos. A la diferencia entre el límite inferior y superior de una clase se le llama frecuencia de clase.

Cuando las clases no tienen límite inferior y superior se les llama abiertas y a las que están acotadas se les denomina cerradas. Con frecuencia la primera y/o última clase de una distribución es abierta.

Frecuencia Acumulada

La primera variación que podemos mencionar es la frecuencia acumulativa, la cual se obtiene mediante adiciones sucesivas de frecuencia en la columna correspondiente.

A continuación, representamos la tabla de frecuencia acumulativa para los datos de tiempo de espera de clientes:

Tiempo de espera Menor que (segundos)	Números acumulativos de clientes
0-20	21
21-40	56
41-60	98
61-80	133
81-100	152
101-120	162
121	172

Los valores de la derecha representan el acumulado o número total de clientes que esperan durante un tiempo menor al indicado por el límite superior de clase especificado en la columna de la izquierda.

Distribución de Frecuencia

Otra variación se obtiene al convertir la tabla de frecuencia de clases (o tabla acumulativa) en una distribución correspondiente, la cual se obtiene como resultado de dividir cada frecuencia de clase (o frecuencia acumulativa) entre el valor total de datos. La distribución de frecuencia se usa particularmente cuando se comparan dos o más distribuciones.

4.5 Representación de las Distribuciones

La frecuencia y distribución acumulativa representan algunas veces en forma gráfica la interpretación de los datos. La más común de las representaciones gráficas es el histograma, el cual representa la frecuencia de clase como rectángulos cuyas longitudes son proporcionarles a las frecuencias de clases.

La primera consideración en la construcción de distribuciones de frecuencia es la especificación del número de clases y los límites mayores e inferiores de cada clase.

Esta elección depende de la naturaleza y uso final de los datos, sin embargo, se recomienda por lo general la siguiente guía:

- Siempre que sea posible, la amplitud de las clases debe ser de igual extensión. Excepto la primera y última clase que por lo general son abiertas.
- Los intervalos de clases no deben solaparse y cada dato debe caer en una clase. En otras palabras, cada dato debe ser asignable a una y solamente una clase.
- Normalmente se usan como mínimas cinco clases y no más de 20.

4.6 El Papel de las Computadoras

El dar sentido a los datos se ha hecho muy fácil con el advenimiento de las computadoras modernas. Algunos métodos requieren tal volumen de cálculos que sería imposible realizarlos sin ellas. Las computadoras pueden hacer gráficas de los datos de manera sencilla y con claridad. Los gerentes tienen abierta la oportunidad de explorar los datos en vez de apoyarse en fórmulas estereotipadas, y es posible comunicar los resultados con claridad y efectividad por medio de gráficas y diagramas. Todo método estadístico presenta hipótesis; con un programa adecuado para la computadora es posible comprobar lo razonable que son las hipótesis.

Literalmente existen cientos de paquetes de software estadístico para los datos. Algunos como Minitab, SAS, SPSS y BMDP fueron desarrollados originalmente para macrocomputadoras y posteriormente adaptados para computadoras personales. Otros, como StatGrafic, Execustat, Systat (para IBM y compatibles) y Data Desk (para Apple Macintosh) fueron escritos para computadoras personales.

4.7 Prueba de Hipótesis

La ciencia va construyendo afirmaciones provisorias sobre cómo son los objetos que estudia y las relaciones que hay entre ellos, estas suposiciones son las **hipótesis.** Sólo son consideradas **hipótesis** aquellas afirmaciones que brindan información precisa sobre la realidad y que pueden ser sometidas a prueba. La hipótesis tiene valor **predictivo**, es decir que, en caso de ser aceptada puede anticipar lo que va a ocurrir.

Las hipótesis proponen tentativamente las respuestas a las preguntas de investigación; la relación entre ambas es directa e íntima. Relevan a los objetivos y preguntas de investigación para guiar el estudio. Por ello, estas surgen de los objetivos y preguntas de investigación, una vez que han sido reevaluadas a raíz de la revisión de la literatura.

Con frecuencia los problemas a los que se enfrenta el ingeniero, el científico, o el administrador, es el tener que formular un procedimiento de decisión basado en los datos que conduzcan a una conclusión acerca de algún planteamiento científico. Esta es una situación en que se encuentra, un investigador que pretende demostrar que la droga A es más efectiva que la droga B, un ingeniero de control de calidad desea comprobar si un proceso de producción origina mejores resultados que el que se encuentra en operación.

En cada uno de los casos el responsable de este estudio postula o conjetura algo acerca del sistema. Esto constituye enunciados provisionales, puesto que al no poder integrar el cumulo de conocimientos concerniente a una situación, aparea la incertidumbre.

Y si hay que tomar decisiones es porque hay alterativas. El proceso de enfrentar las hipótesis al tomar como punto de apoyo los datos muéstrales, constituye lo que llamamos Prueba o Contraste de Hipótesis. Por ejemplo, un ingeniero de control de calidad de una fábrica de tornillos tiene la sospecha que el proceso de producción de tornillos no se está cumpliendo con la especificación promedio que debe ser de 5 cm. Y que por el contrario es menor.

Si esto es así hay que ajustar la producción para que reúna las especificaciones necesarias, y hay dos posibilidades:

> **Hipótesis Nula H_0:** La longitud promedio es de 5 cm. **$H_0 = 5$ cm.**
> **Hipótesis Alterna H_a:** Longitud media de los tornillos es menor a 5 cm. **$Ha < 5$ cm.**

Confrontación: $H_0 = 5$ cm Vs. $Ha < 5$ cm

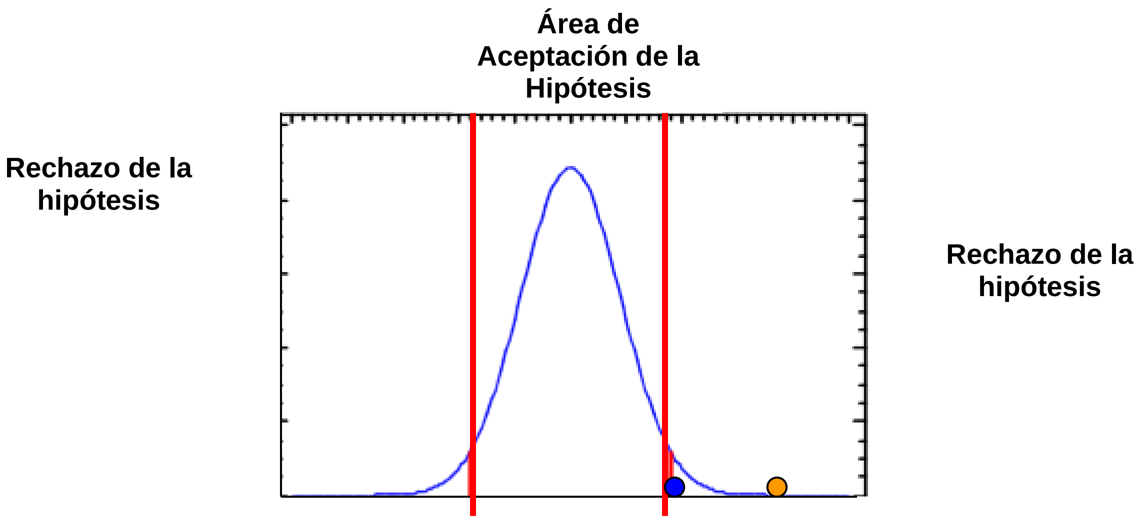
La escritura nos dice que existe una confrontación de afirmación y solo la evidencia de los datos nos podrá indicar hacia donde nos inclinamos, lo que no quiere decir que queda demostrada, sino que no queda inválida ante la evidencia de la muestra.

Hipótesis Nula e Hipótesis Alterna

La Hipótesis Nula en su planteamiento debe expresar la condición de "**no cambio**" de allí el nombre de Hipótesis nula. La hipótesis alterna representa la proposición hipotética que se espera aceptar, su planteamiento debe recoger lo que parece ser más verosímil. En cualquier investigación, encontramos la existencia de dos teorías o hipótesis implícitas, que denominaremos hipótesis nula e hipótesis alternativa, que de alguna manera reflejarán esa idea a priori que tenemos y que pretendemos contrastar con la ``realidad''. La Hipótesis Nula y la Hipótesis Alterna deben ser excluyentes, aunque no siempre.

Una vez identificada la H_o y H_a, se procede a reunir evidencias cuyos resultados apoyen una hipótesis (distribución) o la otra, esta evidencia se resume mediante un valor obtenido por un estadístico de prueba. Calculando este valor y al seguir determinada regla, Regla de Decisión se acepta o se rechaza la Ho.

Para decidir si se rechaza o no la hipótesis nula se ve el valor de un estadístico de prueba de la muestra. La decisión de rechazar se toma si la estadística de prueba cae en una región que llamamos *región crítica* o *región de rechazo*.
Las regiones de rechazo se presentan de acuerdo con la hipótesis alterna, para determinar



la región crítica, tenemos en cuenta la significación que queremos para la prueba, ésta se refleja en el valor crítico que se lee en la tabla normal o t, según sea el caso, o la tabla de puntos porcentuales para el Estadígrafo de Prueba Kolmogorov-Smirnov o la tabla de valores X^2 para cada caso.

De la misma manera aparecen, implícitamente, diferentes tipos de errores que podemos cometer durante el procedimiento. No podemos olvidar que, habitualmente, el estudio y las conclusiones que obtengamos para una población cualquiera, se habrán apoyado exclusivamente en el análisis de sólo una parte de ésta. De la probabilidad con la que estemos dispuestos a asumir estos errores, dependerá, por ejemplo, el tamaño de la muestra requerida.

Ejemplo:

Supongamos que debemos realizar un estudio sobre la altura media de los habitantes de cierto pueblo de España, antes de tomar una muestra, lo lógico es hacer la siguiente suposición a priori, (hipótesis que se desea contrastar y que denotamos H_0)
Al obtener una muestra de tamaño n = 8, podríamos encontrarnos ante uno de los siguientes casos:
1. Muestra = {1,50; 1,52; 1,48; 1,55; 1,60; 1,49; 1,55; 1,63}
2. Muestra = {1,65; 1,73; 1,52; 1,75; 1,65; 1,75; 1,78} 1,80

Intuitivamente, en el caso 1, sería lógico suponer que la muestra obtenida sobre los habitantes del pueblo sea muy poco representativa, la hipótesis H_0 debe ser rechazada, la altura de los habitantes de ese pueblo de España es mayor. En el caso 2 tal vez no podamos afirmar con rotundidad que la hipótesis H_0 sea cierta, sin embargo, no podríamos decidir si es cierta o no, lo que sí sabemos es que algún grado de error existe.

4.7.1 Nivel de Significación

La probabilidad máxima con la que en el ensayo de una hipótesis se puede cometer un Error del Tipo I se llama nivel de significación del ensayo. Esta probabilidad se denota frecuentemente por α; generalmente se fija antes de la extracción de las muestras, de modo que los resultados obtenidos no influyen en la elección de descartarla y la admitimos por una cuestión de simplicidad. Cuando tomamos decisiones referentes a un todo poblacional en base a una parte de este todo, la muestra; nos exponemos a equivocarnos y a cometer errores respecto a la situación verdadera de este todo.

1. Rechazamos la H_0 al ser verdadera **(Error Tipo I)**
2. Aceptamos H_0 al ser falsa **(Error Tipo II)**

4.8 Técnicas Adicionales de Análisis de Datos

Hasta este punto hemos tratado con datos provenientes de mediciones sobre una variable particular que generalmente estaba en una escala continua. Algunos ejemplos son dimensiones, voltajes y dureza Rockwell. En estas situaciones, generalmente se suponía que la distribución fundamental era conocida, salvo posiblemente por parámetros acerca de los cuales había que hacer pruebas e hipótesis u obtener estimados.

En muchas situaciones estos supuestos podrían justificarse por consideraciones teóricas tales como el teorema del límite central o notando que los estadígrafos de prueba o estimadores eran robustos. Un procedimiento decisorio se conoce como robusto si su comportamiento no es sensible a desviaciones de los supuestos fundamentales. Cuando este tipo de información no estaba disponible, se sugerían procedimientos no paramétricos. Sin embargo, se ha dicho poco sobre verificar hechos acerca de la distribución de probabilidad fundamental.

4.8.1 Técnicas Cualitativas para Determinar la Forma de una Distribución

El problema de determinar la forma de una distribución es probablemente uno de los problemas más importantes encontrados en la práctica, pero también uno de los más difíciles. Cuando sea posible el investigador usa procedimientos decisorios o estimadores que son robustos. Sin embargo, hay ocasiones cuando el necesita verificar los supuestos

Acerca de la forma de la distribución fundamental:

1. Construir un histograma

2. Una función de distribución acumulativa muestral y compararlos con alguna función de densidad anticipada o distribución de probabilidad, así la variable es discreta. Dado que esta es una medida cualitativa adolece de incapacidad de distinguir entre anomalías en formas debido a fluctuaciones de muestreo y consideraciones teóricas.
3. Datos representados en papel de probabilidad. Por esta transformación monótona apropiada de la escala vertical de la gráfica de la función de distribución acumulativa, es posible transformar esta curva en una línea recta.

4.8.2 Técnicas Cuantitativas para Determinar la Forma de una Distribución

Las técnicas cualitativas presentadas anteriormente sufren de defecto, ya que las mismas no presentan criterios objetivos para juzgar si los datos se ajustan a las supuestas distribuciones. Se tratarán dos pruebas estadísticas para este propósito:

- La Prueba Kolmogorov-Smirnov.
- Prueba Chi-Cuadrado para la Bondad de Ajuste.

4.8.2.1 La Prueba Kolmogorov-Smirnov

La prueba Kolmogorov-Smirnov, bautizada así en honor de los estadísticos A.N. Kolmogorov y N.M. Smirnov quienes la desarrollaron, se trata de un método no paramétrico sencillo para probar si existe una diferencia significativa entre una distribución de frecuencia observada, y otra teórica. La Prueba Kolmogorov-Smirnov es, por consiguiente, otra medida de la bondad de ajuste de la distribución de frecuencia teórica, como lo es la de Chi-Cuadrado. Sin embargo, la Prueba Kolmogorov-Smirnov tiene varias ventajas sobre la Prueba Chi- Cuadrado, es más poderosa y es más fácil de usar puesto que no requiere que se agrupen los datos de alguna manera.

¿En qué consiste la Prueba Kolmogorov-Smirnov?

La prueba se fundamenta en la diferencia absoluta máxima D entre los valores de la distribución acumulada de una muestra aleatoria de tamaño n y la distribución teórica determinada. Para decidir si esta diferencia es mayor de lo que razonablemente puede esperarse con un nivel de significancia determinada, buscamos los valores críticos de D en una tabla de valores críticos.

La prueba Kolmogorov-Smirnov sirve para encontrar el grado de confianza con que se puede afirmar que un conjunto de datos sigue un comportamiento semejante al que se propone como representativo. Este comportamiento propuesto frecuentemente se representa por la ecuación que describe la distribución que, se presume, tienen los datos.

Cálculo de la Estadística de la Prueba Kolmogorov-Smirnov

Desviación absoluta máxima de F_e desde F_o. Si

Dn = max	F_e - F_o	

la hipótesis de que la frecuencia esperada no es igual a la frecuencia teórica cuando Dn > D_α, donde d_α es el valor dado en la tabla A, es el criterio decisivo para la aceptación o el rechazo.

La prueba de **Kolmogorov-Smirnov** sirve para ratificar si la distribución de probabilidad sugerida para un sistema es la adecuada. Se implementará cumpliendo con los siguientes pasos:

- Basándose en los datos recopilados y en el histograma de frecuencias se plantea una hipótesis sobre la posible distribución de probabilidad de los datos recopilados.
- Se calcula la probabilidad de cada clase de frecuencia con los datos recogidos, al igual que las probabilidades basándose en la fórmula de la distribución de probabilidad propuesta.
- La desviación absoluta se calcula restando las columnas de las probabilidades acumuladas.
- Con esta diferencia se ubica una columna en la tabla de Kolmogorov-Smirnov. En esta columna y dependiendo del valor de la muestra n se utilizará un valor definido en la fila correspondiente a n, en caso de ser n < 35 o una fórmula n > 35.

Por último, se aprueba o se rechaza la hipótesis sugerida dependiendo del valor encontrado en la tabla/fórmula. Si este resulta mayor o igual a la desviación absoluta obtenida en el paso 3, la hipótesis se aprueba; en caso contrario se rechaza.

Si el valor no se encuentra en la tabla de puntos porcentuales para el Estadígrafo Kolmogorov- Smirnov se utiliza la siguiente fórmula:

α	**20%**	**15%**	**10%**	**5%**	**1%**
Fórmula aproximada para N muy grandes	$1.07//\sqrt{n}$	$1.14//\sqrt{n}$	$1.22//\sqrt{n}$	$1.36//\sqrt{n}$	$1.63//\sqrt{n}$

Tabla A
Valores Críticos de la Prueba de Kolmogorov – Smirnov

Tabla de los Puntos Porcentuales para el Estadígrafo de Prueba Kolmogorov – Smirnov

Tamaño de la muestra N Grados de Libertad	Nivel de Significancia				
	20%	15%	10%	5%	1%
1	0.9	0.925	0.950	00975	0.995
2	0.84	0.726	0.776	0.842	0.929
3	0.565	0.597	0.642	0.708	0.829
4	0.494	0.525	0.564	0.624	0.734
5	0.446	0.474	0.510	0.563	0.669
6	0.410	0.436	0.470	0.521	0.618
7	0.381	0.405	0.438	0.486	0.577
8	0.358	0.381	0.411	0.457	0.543
9	0.339	0.360	0.388	0.432	0.514
10	0.322	0.342	0.368	0.409	0.486
11			0.352	0.391	0.468
12			0.338	0.375	0.450
13			0.325	0.361	0.433
14			0.314	0.349	0.418
15			0.304	0.338	0.404
16			0.295	0.328	0.392
17			0.286	0.318	0.381
18			0.278	0.309	0.371
19			0.272	0.301	0.363
20			0.264	0.294	0.352
25			0.24	0.264	0.317
30			0.22	0.242	0.290
35			0.21	0.23	0.27
40				0.210	0.252
50				0.188	0.226
60				0.172	0.207
70				0.160	0.192
80				0.150	0.180
90				0.141	
100				0.134	

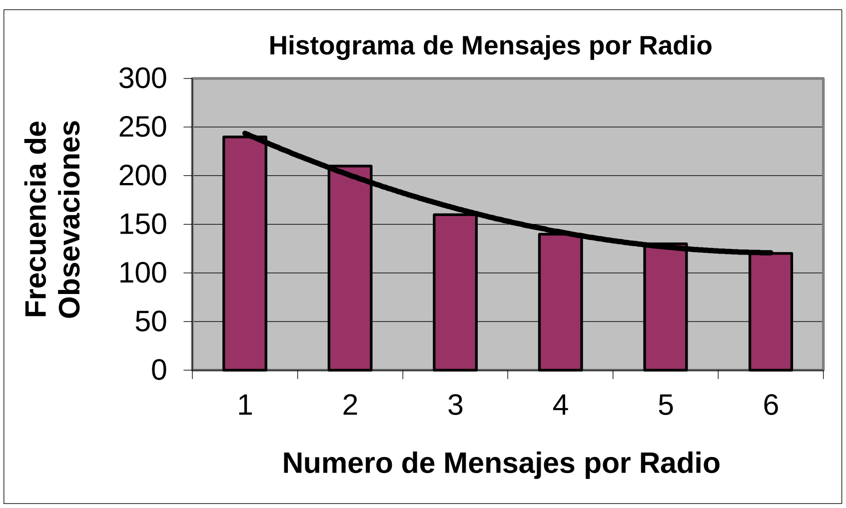
Ejemplo 1

Tomamos como base los datos de la tabla de distribución de frecuencia de consultas telefónicas por intervalo de una hora, descrita en la parte inferior, para ajustarla a una distribución de Poisson a un nivel de significancia de error α **= 0.05**

Tabla de Distribución de Frecuencias

Número de consultas (X)	**Número de intervalos en una 1 hora con X consultas** (Frecuencia observada)
0	240
1	140
2	70
3	45
4	25
5	3
	Σ= 523

Histograma de frecuencia por radio



Paso 1: Recolectar datos, es decir, crear nuestra muestra.

Paso 2: Crear la distribución de frecuencia, como se muestra en la tabla anterior.

Paso 3: Plantear la hipótesis.

H_0: No existe diferencia significativa entre los datos observados(empíricos) y aquellos que serán producidos. Teóricamente por una distribución de Poisson.
H_a: Existe diferencia significativa entre los datos observados (empíricos) y aquellos que serán producidos teóricamente. Por una distribución Poisson.

Paso 4: Calcular la probabilidad observada.

La probabilidad observada se calcula dividiendo cada frecuencia observada entre el total de frecuencia, es decir, por ejemplo:

$$\frac{240}{523} = 0.458$$

y así sucesivamente hasta obtener la siguiente tabla:

# de consultas (X)	# de intervalos en una 1 hora con X consultas (Frecuencia observada)	Probabilidad observada
0	240	0.459
1	140	0.268
2	70	0.134
3	45	0.086
4	25	0.048
5	3	0.005
	Σ= 523	Σ = 1.00

Paso 5: Calcular los parámetros de la muestra.

Para este ejemplo, calcular los parámetros que se utilizarán en la fórmula de distribución de Poisson, como λ. Para calcular λ, se debe calcular la media y la varianza de la distribución de frecuencia, o sea:

Mi Fi = multiplicar cada valor de la frecuencia por su correspondiente valor de x.

Mi^2Fi = elevar cada valor de X al cuadrado y luego multiplicarlo por su correspondiente frecuencia.

Mi Valores de x	Fi Frecuencia	Mi Fi	Mi ^{2} Fi
0	240	0	0
1	140	140	140
2	70	140	280
3	45	135	405
4	25	100	400
5	3	15	75
	Σ= 523	**Σ= 530**	**Σ= 1300**

Fórmula de la media:

$$\bar{X} = \frac{\sum_{i=1}^{k} Mi\,Fi}{n} = \frac{530}{523} = 1.01$$

Fórmula de la varianza:

$$\bar{S} = \frac{\sum_{i=1}^{k} Mi^2\,Fi - n\,\bar{x}^2}{n-1} = \frac{1300 - 523\,(1.01)^2}{523-1} = 1.46$$

Entonces una vez obtenidos los valores de la media y la varianza procedemos a calcular λ.

$$\lambda = \frac{X + S}{2} \quad \frac{1.01 + 1.46}{2} - 1.23$$

Paso 6: Calcular la probabilidad teórica, utilizando la fórmula de la probabilidad planteada en el paso 3.

Para este caso se utiliza la fórmula de **Poisson** para calcular la probabilidad teórica:

$$P(x) = \frac{\lambda^x\, e^{-\lambda}}{X!}$$

Una vez calculado todos los valores, la tabla resultante es la siguiente:

Número de consultas (X)	**Número de intervalos en 1 hora con N consultas (Frecuencia observada)**	**Probabilidad observada**	**Probabilidad teórica**
0	240	0.459	0.29
1	140	0.268	0.36
2	70	0.134	0.22
3	45	0.086	0.09
4	25	0.048	0.03
5	3	0.005	0.01
	Σ= 523	**Σ = 1.00**	**Σ = 1.00**

Paso 7: Calcular la probabilidad acumulada de la probabilidad observada. Esta se calcula asignando el primer valor de la probabilidad observada, como primer valor de la probabilidad acumulada de la probabilidad observada. Una vez hecho esto, los valores restantes se obtienen sumando el último valor obtenido de la probabilidad acumulada de la probabilidad observada al siguiente valor de la probabilidad observada.

# de consultas (X)	**# de intervalos en una 1 hora con X consultas (Frecuencia observada)**	**Probabilidad observada**	**Probabilidad teórica**	**Probabilidad acumulada de la probabilidad observada**
0	240	0.459	0.29	0.459
1	140	0.268	0.36	0.727
2	70	0.134	0.22	0.861
3	45	0.086	0.09	0.947
4	25	0.048	0.03	0.995
5	3	0.005	0.01	1.00
	Σ= 523	**Σ = 1.00**	**Σ = 1.00**	

Paso 8: Calcular la probabilidad acumulada de la probabilidad teórica, esta se calcula de la misma manera que la probabilidad acumulada de la probabilidad observada, la única diferencia es que se utilizan los datos de la columna de la probabilidad teórica.

Una vez realizados los cálculos, se obtiene la siguiente tabla:

# de consultas (X)	# de intervalos en una 1 hora con X consultas (Frecuencia observada)	Probabilidad observada	Probabilidad teórica	Probabilidad acumulada de la probabilidad observada	Probabilidad acumulada de la probabilidad teórica
0	240	0.459	0.29	0.459	0.29
1	140	0.268	0.36	0.727	0.65
2	70	0.134	0.22	0.861	0.87
3	45	0.086	0.09	0.947	0.96
4	25	0.048	0.03	0.995	0.99
5	3	0.005	0.01	1.00	1.00
	Σ= 523	**Σ = 1.00**	**Σ = 1.00**		

Paso 9: Calcular el discriminante calculado, que se obtiene restando la probabilidad acumulada de la probabilidad observada de la probabilidad acumulada de la probabilidad teórica. Se obtiene la siguiente tabla:

# de consultas (X)	# de intervalos en una 1 hora con X consultas (Fo)	Probabilidad observada	Probabilidad teórica (Fe)	Probabilidad acumulada de la probabilidad observada	Probabilidad acumulada de la probabilidad teórica	Discriminante calculado
0	240	0.459	0.29	0.459	0.29	0.17
1	140	0.268	0.36	0.727	0.65	0.08
2	70	0.134	0.22	0.861	0.87	0.01
3	45	0.086	0.09	0.947	0.96	0.01
4	25	0.048	0.03	0.995	0.99	0.00
5	3	0.005	0.01	1.00	1.00	0.00
	Σ= 523	**Σ = 1.00**	**Σ = 1.00**			

Paso 10: Seleccionar el mayor de los valores del discriminante calculado. De los valores que se obtienen se escoge el mayor de ellos en este caso el discriminante calculado será igual a **0.17**.

Paso 11: Calcular el discriminante crítico. Este se obtiene empleando la tabla de **Kolmogorov-Smirnov**, utilizando el tamaño de muestra y el nivel de significancia de error.

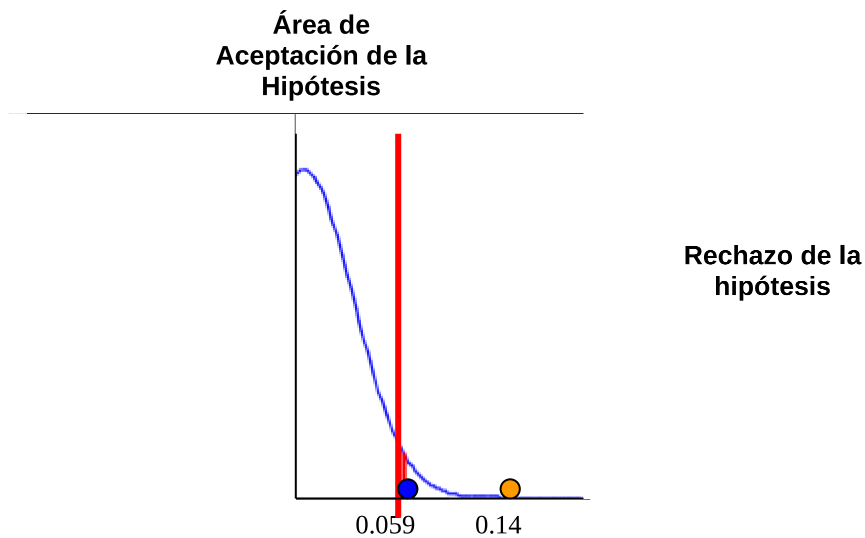
$$D_{\text{crítico}} = \frac{1.36}{(n)^{1/2}} = \mathbf{0.059}$$

Paso 12: Comparar el discriminante crítico contra el discriminante calculado.

Paso 13: Si el discriminante crítico es mayor que el discriminante calculado, entonces se acepta la hipótesis, de lo contrario se rechaza.

Conclusión:

Se rechaza la hipótesis $\mathbf{H_0}$, ya que el discriminante crítico es menor que el discriminante calculado. Y la distribución observada no es representativa a la distribución de Poisson.

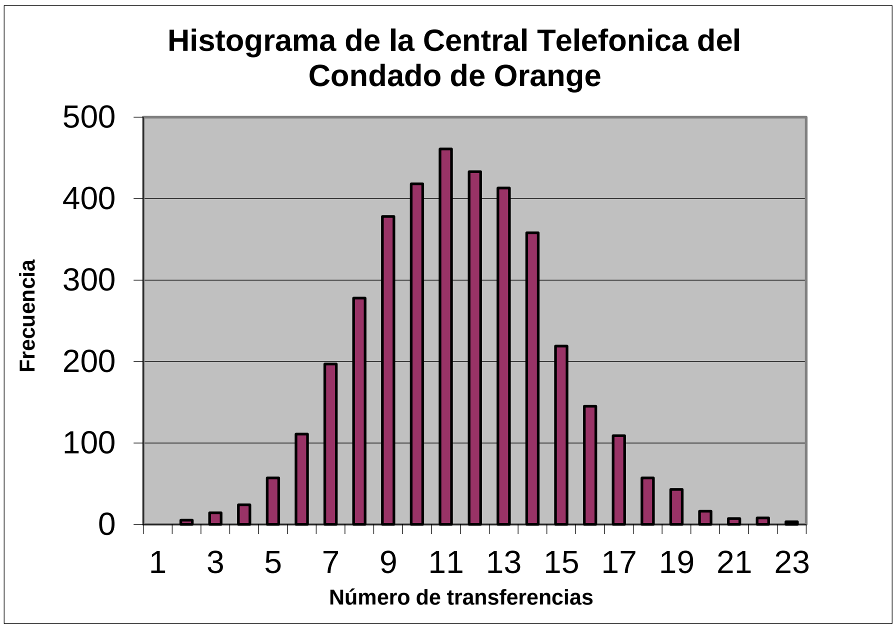


Ejemplo 2

Supongamos que la central telefónica del Condado de Orange, en California, ha estado llevando un registro del número de transacciones usado en un instante dado. Fueron 3754 observaciones en ocasiones distintas. Para propósito de inversión de capital, el funcionario de presupuesto de la compañía piensa que el patrón de uso sigue una distribución de Poisson con una media de 8.5. Se desea probar esta hipótesis al nivel de significancia de 0.01.

Solución

Paso 1: Recolectar datos, es decir, crear nuestra muestra.



Paso 2: Crear la distribución de frecuencia, como se muestra en la tabla.

Número de transacciones	Frecuencia Observada
0	0
1	5
2	14
3	24
4	57
5	111
6	197
7	278
8	378
9	418
10	461
11	433
12	413
13	358
14	219
15	145
16	109
17	57
18	43
19	16
20	7
21	8
22	3
	Σ =3754

Paso 3: Plantear la hipótesis

Formulación de las Hipótesis

H_o: una distribución de Poisson con λ = 8.5 es una buena descripción del patrón de uso.
H_a: una distribución de Poisson con λ = 8.5 no es una buena descripción del patrón de uso.

Los datos enumerados se representan en la tabla a continuación, ahora podremos utilizar la fórmula de Poisson para calcular la frecuencia esperada.

Paso 4: Calcular la probabilidad observada.

La probabilidad observada se calcula dividiendo cada frecuencia observada entre el total de frecuencia.

Número de transacción	Frecuencia observada	Frecuencia acumulativa observada Fo acumulada n	Frecuencia acumulativa observada
0	0	0	0.0000
1	5	5	0.0013
2	14	19	0.0051
3	24	43	0.0115
4	57	100	0.0266
5	111	211	0.0562
6	197	408	0.1087
7	278	686	0.1827
8	378	1064	0.2834
9	418	1482	0.3948
10	461	1943	0.5176
11	433	2376	0.6329
12	413	2789	0.7429
13	358	3147	0.8383
14	219	3366	0.8966
15	145	3511	0.9353
16	109	3620	0.9643
17	57	3677	0.9795
18	43	3720	0.9909
19	16	3736	0.9952
20	7	3743	0.971
21	8	3751	0.9992
22	3	3754	1
	Σ =3754		

Paso 5: Calcular la probabilidad teórica, utilizando la fórmula de la probabilidad planteada en el paso 3. Para este problema se usa el Estadígrafo de Poisson.

$$P(x) = \frac{\lambda^x \; e^{-\lambda}}{X!}$$

Probabilidad Frecuencia Esperada Acumulada o Teórica

Número de transacciones	Frecuencia observada	Frecuencia acumulativa esperada Fe acumulada N	Frecuencia Acumulativa observada	Frecuencia acumulativa observada esperada
0	0	0	0.0000	0.002
1	5	5	0.0013	0.0019
2	14	19	0.0051	0.0093
3	24	43	0.0115	0.0301
4	57	100	0.0266	0.0744
5	111	211	0.0562	0.1496
6	197	408	0.1087	0.2562
7	278	686	0.1827	0.3856
8	378	1064	0.2834	0.5231
9	418	1482	0.3948	0.6530
10	461	1943	0.5176	0.7634
11	433	2376	0.6329	0.8487
12	413	2789	0.7429	0.9091
13	358	3147	0.8383	0.9486
14	219	3366	0.8966	0.9726
15	145	3511	0.9353	0.9862
16	109	3620	0.9643	0.9934
17	57	3677	0.9795	0.9970
18	43	3720	0.9909	0.9987
19	16	3736	0.9952	0.9995
20	7	3743	0.971	0.9998
21	8	3751	0.9992	0.999
22	3	3754	1	1
	Σ =3754			

Paso 6: Calcular el discriminante.

Se obtiene restando la probabilidad acumulada de la probabilidad observada de la probabilidad acumulada de la probabilidad teórica, de lo que nos interesa solo es el valor absoluto.

$$\textbf{Discriminante}_n = \textbf{max} \quad \textbf{Prob. Acum. Teórica – Prob Acum. Observada}$$

Diferencia Absoluta entre las Probabilidades Acumuladas de la Frecuencia Observadas y las Frecuencias Teóricas.

Número de transacciones	Frecuencia observada	Frecuencia acumulativa observada Fe acumulada N	Frecuencia acumulativa observada	Frecuencia acumulativa observada esperada	Fe -Fo
0	0	0	0.0000	0.002	0.0002
1	5	5	0.0013	0.0019	0.0006
2	14	19	0.0051	0.0093	0.0042
3	24	43	0.0115	0.0301	0.0186
4	57	100	0.0266	0.0744	0.0478
5	111	211	0.0562	0.1496	0.0934
6	197	408	0.1087	0.2562	0.1475
7	278	686	0.1827	0.3856	0.2029
8	378	1064	0.2834	0.5231	0.2397
9	418	1482	0.3948	0.6530	0.2582
10	461	1943	0.5176	0.7634	0.2458
11	433	2376	0.6329	0.8487	0.2158
12	413	2789	0.7429	0.9091	0.1662
13	358	3147	0.8383	0.9486	0.1103
14	219	3366	0.8966	0.9726	0.0760
15	145	3511	0.9353	0.9862	0.0509
16	109	3620	0.9643	0.9934	0.0291
17	57	3677	0.9795	0.9970	0.0175
18	43	3720	0.9909	0.9987	0.0078
19	16	3736	0.9952	0.9995	0.0043
20	7	3743	0.971	0.9998	0.0027
21	8	3751	0.9992	0.999	0.0007
22	3	3754	1	1	0
	Σ =3754				

Al comparar estas frecuencias esperadas con las frecuencias observadas, podremos examinar el alcance de la diferencia entre ellas: la desviación absoluta. En la tabla las frecuencias acumulativas absolutas F_0, frecuencia acumulativa esperada F_e y la desviación absoluta para X= 0 – 22.

Paso 7: Seleccionar el mayor de los valores del discriminante calculado.
De los valores que se obtienen se escoge el mayor de ellos en este caso el discriminante calculado será igual a **0.2582**.

Paso 8: Calcular el discriminante crítico.

Este se obtiene empleando la tabla de **Kolmogorov-Smirnov**, utilizando el tamaño de muestra y el nivel de significancia de error.

La estadística Kolmogorov - Smirnov debe ser una prueba de extremos. Los valores críticos para Dn se han tabulado en la tabla A. Si se busca en la fila n = 3754 el tamaño de la muestra y la columna en el nivel de significancia de 0.01 encontramos que el valor crítico de Dn debe calcularse usando a la fórmula:

$$D_n = \frac{1.63}{(n)^{1/2}} = \mathbf{0.0266}$$

Paso 9: Comparar el discriminante crítico contra el discriminante calculado.

Paso 10: Si el discriminante crítico es mayor que el discriminante calculado, entonces se acepta la hipótesis, de lo contrario se rechaza.

Conclusión:

El siguiente paso es comparar el valor calculado de Dn con el valor Critico de Dn que se encuentra en la tabla. Si el valor de la tabla para el nivel de significancia elegido es mayor que el valor calculado de Dn, entonces aceptamos la hipótesis nula. Obviamente 0.0266 < 0.2582, así que rechazamos H_o, y llegamos a la conclusión de que una distribución de Poisson con una media 8.5, no es una buena descripción del patrón de uso del transmisor en la central telefónica del condado de Orange.

4.8.2.2 Teoría de Bondad de Ajuste

Prueba Chi Cuadrado para la prueba de la Bondad de Ajuste.

La prueba Chi-Cuadrado puede usarse también para decidir si una distribución en particular, como lo es Binomial, Poisson, o Normal, es la distribución apropiada. Esta es una habilidad importante porque como tomadores de decisiones que usamos las estadísticas, necesitaremos escoger una cierta distribución de probabilidad para representar la distribución de los datos que tengamos que trabajar.

Necesitamos saber que tan lejos podemos avanzar con la suposición que subyacen en una distribución particular antes de que debamos concluir que tal distribución ya no se puede aplicar. La prueba Chi-Cuadrado nos permite hacer esta pregunta y probar si existe una diferencia significativa entre una distribución de frecuencia observada y una distribución de frecuencia teórica. De esta manera podemos determinar la bondad del ajuste de una distribución teórica (es decir, que también se pueden ajustar los datos que observamos a la realidad). Así pues, podemos determinar si debemos creer que los datos observados constituyen una muestra obtenida de la distribución teórica hipotetizada.

Hablamos de bondad de ajuste cuando tratamos de comparar una distribución de frecuencia observada con valores correspondientes de una distribución esperada o teórica. Estadígrafo de prueba Chi-Cuadrado para la bondad de ajuste:

$$X^2 = \frac{(Fo - fe)^2}{Fe}$$

La distribución muestral de este Estadígrafo es aproximada por la distribución Chi-Cuadrado con K-m grados de libertad donde k es el número de términos en la fórmula, aunque se calcula X^2 y m es el número de cantidades, conseguidas de los datos observados, que se necesitan para calcular frecuencias observadas.

En donde F_o es la frecuencia observada o empírica y la F_e son los datos teóricos. Los datos deben ser agrupados de tal manera que la frecuencia esperada sea mayor que 5.

Tabla de Valores X^2_α

Tamaño de la muestra N	Nivel de Significancia				
	.20	.10	.05	.025	.01
1	1.642	2.706	3.841	5.024	6.635
2	3.219	4.605	5.991	7.378	9.210
3	4.642	6.251	7.815	9.348	11.345
4	5.989	7.779	9.488	11.143	13.277
5	7.289	9.236	11.07	12.832	15.056
6	8.558	10.645	12.592	14.449	16.812
7	9.803		14.067	16.013	18.475
8	11.030		15.507	17.535	20.090
9	12.242		16.919	19.023	12.666
10	13.442		18.307	20.483	23.209
11	14.631		19.675	21.920	24.725
12	15.812		21.026	23.337	26.217
13	16.985		22.362	24.736	27.688
14	18.151		23.685	27.488	29.141
15			24.996	28.845	30.578
16			26.296	30.191	32.000
17			28.869	31.526	33.409

Ejemplo 1

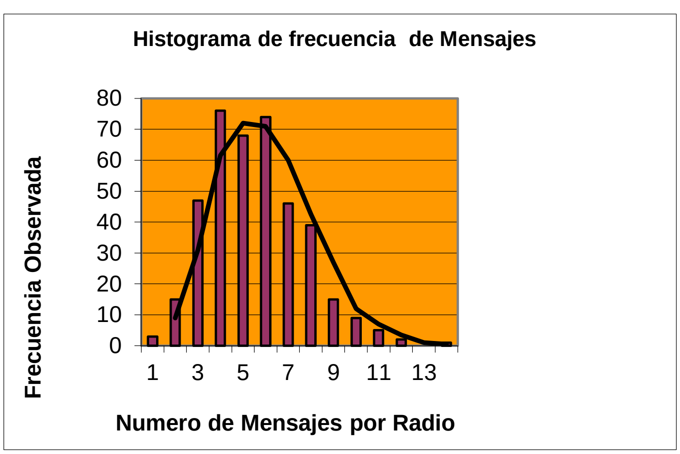
Supóngase que durante 400 intervalos de 5 minutos el control central de tráfico aéreo de un aeropuerto recibió 0, 1, 2, 3, 4, 5, 6, 7, 8, 9,.. .12, 13 mensajes de radio con frecuencias respectivas de 3, 15, 47, 76, 68, 74, 46, 39, 15, 9, 5, 2, 0, 1. Supóngase además que queremos verificar si estos datos apoyan la afirmación de que el número de mensajes de radio que reciben durante un intervalo de 5 minutos puede considerarse una variable aleatoria que tiene una Distribución Poisson con, λ= 4.6

Paso 1: Recopilación de Datos

Tabla de frecuencia observada y número de mensajes por radio

Número de Mensajes por Radio	Frecuencia Observada	
0	3	18
1	15	
2	47	
3	76	
4	68	
5	74	
6	46	
7	39	
8	15	
9	9	
10	5	8
11	2	
12	0	
13	1	
	Σ= 400	

Paso 2: Crear distribución de frecuencia (Histograma).



Paso 3: Formulación de las Hipótesis

H_0: La variable aleatoria tiene una distribución Poisson con, λ= 4.6, con un Nivel de Significancia de α = 0.01

H_a: La variable aleatoria no tiene distribución Poisson con, λ= 4.6, con un Nivel de Significancia de α = 0.01

Paso 4: Cálculo de la Probabilidad Esperada

Se utiliza el Estadígrafo de Probabilidad, en este caso λ= 4.6 y da como resultado la tabla que a continuación se presenta.

Cálculo de la probabilidad esperada como base para el cálculo de la probabilidad teórica

Número de Mensajes por Radio	Frecuencia Observada	Probabilidad de Poisson
0	3	0.010
1	15	0.046
2	47	0.107
3	76	0.163
4	68	0.187
5	74	0.173
6	46	0.132
7	39	0.087
8	15	0.050
9	9	0.025
10	5	0.012
11	2	0.005
12	0	0.002
13	1	0.001
	Σ= 400	**Σ= 1**

Paso 5: Cálculo de la Probabilidad Teórica

La probabilidad de Poisson debe ser multiplicada por la sumatoria de observaciones 400; y da como resultado la frecuencia esperada**.**

Frecuencia Real y Teórica Esperada

Número de Mensajes por Radio	**Frecuencia Observada**	**Probabilidad de Poisson**	**Frecuencia Esperada**
0	3	0.010	4
1	15	0.046	18.4
2	47	0.107	4208
3	76	0.163	65.2
4	68	0.187	74.8
5	74	0.173	69.2
6	46	0.132	52.8
7	39	0.087	34.8
8	15	0.050	20
9	9	0.025	10
10	5	0.012	4.8
11	2 8	0.005	2 8
12	0	0.002	0.8
13	1	0.001	0.4
	Σ= 400	**Σ= 1**	**Σ= 400**

Nótese que hemos combinado algunos de los datos en tal forma que ninguna de las frecuencias esperada sea menor de 5. Para probar si existe discrepancia entre las frecuencias observadas y las esperadas pueden atribuir al alzar.

Paso 6: Diferencia entre la Frecuencia Teórica y la Frecuencia Observable

Se compara la frecuencia observada con la frecuencia esperada. Teniendo la frecuencia observada calculada cuando sacamos la frecuencia esperada tenemos que multiplicar la probabilidad teórica de la variable aleatoria Poisson con la sumaria de todas las frecuencias 400.

Cálculo Estadístico de la Chi-Cuadrado

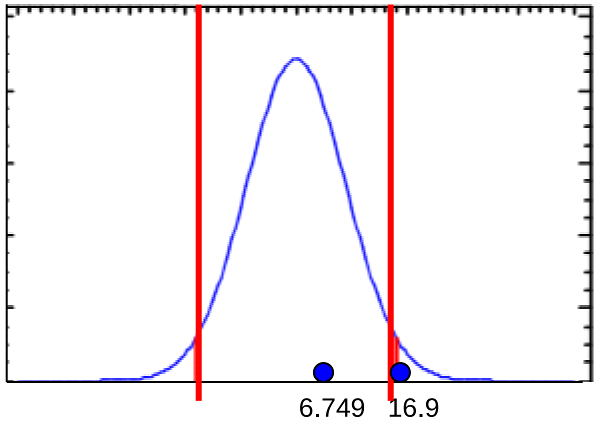
Frecuencia Observada (Fo)	Frecuencia Esperada (fe)	fe- fo	$\frac{(fe-fo)^2}{fe}$
3	4	-1	0.25
15	18.4	-3.4	0.628
47	42.8	4.2	0.412
76	65.2	10.8	1.789
68	74.8	-6.8	0.618
74	69.2	4.8	0.3329
46	52.8	-6.8	0.875
39	34.8	4.2	0.507
15	20	-5	1.25
9	10	-1	0.1
5	4.8	0.2	0.008
2 8	2 8		
0	0.8	0	0
1	0.4		
Σ= 400	**Σ= 400**		**Σ= 6.749**

Paso 8: Utilización de la Tabla

Una vez obtenidos los grados de libertad se proceden a encontrar el nivel de significancia o probabilidad de error, dicha tabla se comporta matricialmente de acuerdo con el obtenido en los cálculos anteriores. Se toma una decisión, o sea la sumatoria de cuadrados de las diferencias de las probabilidades observadas y empíricas**.** Dependiendo si es mayor o menor que el valor de la tabla se concluye, el valor de $X^2_{0.01}$ para K= 10-1= 9 dada solo una cantidad, la frecuencia total de los 400, es necesario en los datos observados para calcular las frecuencias esperadas.

Paso 9: Conclusión

Dado que X^2 no sobrepasa a 16.919, la Hipótesis Nula no puede ser rechazada; se concluye que la distribución de Poisson con $\lambda = 4.6$ es un buen ajuste con $\mathbf{x^2=6.749}$**.**



Ejemplo 2

Utilizaremos la prueba Chi-cuadrada para decidir si aceptamos o no una hipótesis nula que es una hipótesis independiente entre dos variables. Estas dos variables son actitudes hacia la evaluación del desempeño en el trabajo y región geográfica.

Cálculo de las Frecuencias Observadas y Esperadas

Supóngase que una compañía requiere que los recién graduados de la universidad que buscan una colocación en la empresa, sean entrevistados por tres ejecutivos diferentes. Esto permite a la compañía obtener una evaluación condensada de los candidatos. Cada uno de los ejecutivos califica a los candidatos en forma positiva o negativa.

Paso 1: Recolectar los Datos

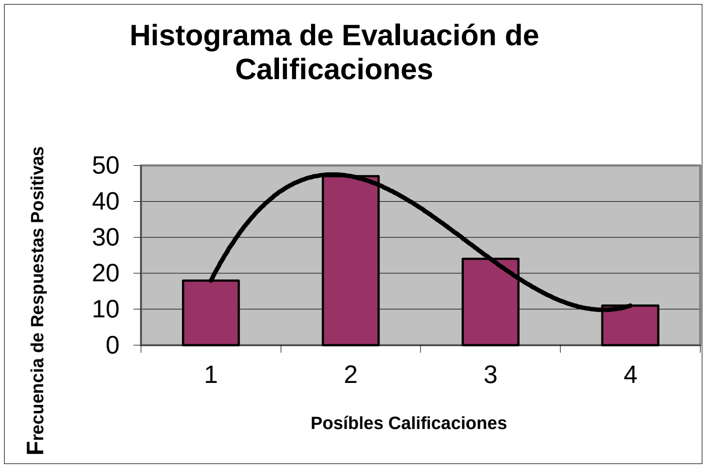
Resultados de las 100 entrevista por los ejecutivos

Posibles calificaciones positivas en las tres entrevistas	Probabilidades binomiales para los resultados
0	18
1	47
2	24
3	11
	Σ= 100

Con el propósito de planificar su fuerza de trabajo, el director de contratación de personal de la compañía piensa que el proceso de entrevista puede ser aproximado por una distribución Binomial con p = 0.40, es decir con una probabilidad de 40% de que cualquier candidato obtendrá una calificación positiva en cualquiera de las entrevistas. Si el director desea probar esta hipótesis a un nivel de significancia de 0.10 **¿De qué manera deberá proceder?**

Paso 2: Crear una distribución de frecuencias como se muestra a continuación.

Planteamiento del problema de manera simbólica



Paso 3: Plantear las Hipótesis

H_o: una distribución Binomial con p = 0.40 es una buena descripción del proceso de entrevistas. **Hipótesis nula.**
H_a: una distribución Binomial con p = 0.40 no es una buena descripción del proceso de entrevistas. **Hipótesis alterna.**

Paso 4: Cálculo de las Probabilidades Teóricas

Cálculo de las probabilidades binomiales

$$P(r) = \frac{n!}{r!\,(n-r)!} * p^r q^{(n-r)}$$

Para resolver el problema debemos determinar si las discrepancias entre las frecuencias observadas y las frecuencias esperadas (si la distribución Binomial fuera el modelo apropiado a utilizar) deberán atribuirse al azar. Podemos determinar cuáles serán las probabilidades binomiales para esta situación en particular. Para las 3 entrevistas, encontraremos la probabilidad de éxito en la tabla de distribución normal buscando la columna etiquetada con n= 3 P= 0.40.

Probabilidades Binomiales para el Problema de las Entrevistas

Posibles calificaciones positivas en las tres entrevistas	Probabilidades binomiales para esos resultados
0	0.2160
1	0.4320
2	0.2880
3	0.0640
	Σ =1

Si se prefiere desistir de la fórmula existe una tabla, esta queda como investigación para el lector, para exactitud se recomienda la fórmula. Teniendo la probabilidad binomial teórica podremos calcular la probabilidad esperada.

Paso 5: Cálculo de las Probabilidades Esperadas

Cálculo de Probabilidades Esperadas

Posibles calificaciones positivas en las tres entrevistas	Probabilidades binomiales para los resultados	Probabilidades binomiales para esos resultados	Número de candidatos entrevistados	Frecuencia esperada de candidatos que obtienen estas calificaciones
0	18	0.2160	100	21.6
1	47	0.4320	100	43.2
2	24	0.2880	100	28.8
3	11	0.0640	100	6.4
	100	**1**	**100**	**100**

Paso 6: Cálculo de la Estadística Chi-Cuadrada

$$X^2 = \frac{(Fo - fe)^2}{Fe}$$

Cálculo de la estadística X^2 de los datos de la entrevista:

Posibles calificaciones positivas en las tres entrevistas	Frecuencia observada	Frecuencia esperada	Fo- fe	$(Fo-fe)^2$	$\frac{(Fo-fe)^2}{fe}$
0	18	21.6	-3.6	12.96	0.6
1	47	43.2	3.8	14.44	0.3343
2	24	28.8	-4.8	23.04	0.8
3	11	6.4	4.6	21.16	3.3063
	100	100			5.406

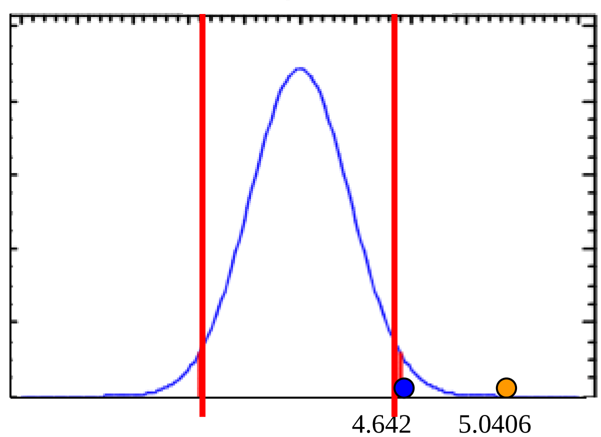
Al comparar la Fe y Fo usando la prueba Chi-Cuadrado, podremos examinar la magnitud de la diferencia entre ellas. Al comparar los resultados de la sumaria de la X^2 y los valores de la tabla podremos concluir si se rechaza o se aprueba.

Paso 7: Conclusión

El valor de Chi-Cuadrado de la muestra es 5.040, cae fuera de la región de aceptación de la hipótesis nula. Por consiguiente, rechazamos la H_0 y llegamos a la conclusión de que la Distribución Binomial con p = 0.40 no logra proporcionar una buena descripción de la frecuencia observada.

Área de Aceptación de la Hipótesis

Rechazo de la hipótesis



Rechazo de la hipótesis

4.9 Bibliografía

DATANALISIS, “Metodologías - Técnicas de Recopilación de la Información”, http://www.datanalisis.com/metodologia_t_recoleccion.htm

Julio Cabrero García, Miguel Richard Martínez, “Recogida de Datos”, http://perso.wanadoo.es/aniorte_nic/apunt_metod_investigac4_9.htm, 18/8/01

Marcelino Auccasi Rojas, “Recolección de Datos”, http://usuarios.lycos.es/enfermeriaperu/investigacion/recodatosprocinfor.htm

Kolmogorov-Smirnov Goodness of Fit Test http://www.itl.nist.gov/div898/handbook/eda/section3/eda35g.htm

Kolmogorov-Smirnov Test http://www.physics.csbsju.edu/stats/KS-test.html