

# JUSTIFICAR LA CONVERSIÓN DE VARIABLES TEÓRICAS DICOTÓMICAS EN VARIABLES OPERATIVAS USANDO LA FUNCIÓN DE DISTRIBUCIÓN BINOMIAL.

AUTOR: *Carlos Javier Gómez Calvo*<sup>1</sup>; *María Consolación Rodríguez Gutiérrez*<sup>2</sup>; *Roberto Fernando Marín Blasco*<sup>2</sup>; *Isabel Martínez Pascual*<sup>1</sup>; *María Edén Lizaranzu Olarte*<sup>1</sup>; *Aitziber Díaz de Cerio Ramsden*<sup>1</sup>

CENTRO DE TRABAJO: 1 Zona Básica de Salud de Nájera. C) Avenida de La Rioja, 5. 26300 Nájera, La Rioja (España). 2 Hospital San Pedro. C) Piqueras, 98. 26006 Logroño, La Rioja (España).

## INTRODUCCIÓN

Objetivo principal:

Emplear la distribución de probabilidad binomial en investigaciones de problemas de salud.

Metodología:

Distribución binomial (función de probabilidad): definición, cálculo. Se simuló un ejemplo pedagógico como apoyo didáctico: en diseños longitudinales, influencia de una variable, “tendencia de ingestión de levotiroxina (ayunas o con alimentos)” en la estabilidad del control analítico de TSH. Aprovechando la distribución binomial, operativización de una variable teórica, “valoración del control analítico de TSH (óptima o subóptima)”, en una variable operativa: valoración óptima (porcentaje de analíticas con resultado normal de TSH  $\geq 90\%$ ) o subóptima ( $< 90\%$ ).

Resultados principales:

Distribución binomial: justifica científicamente la variable operativa, posibilitando cruzar dos variables dicotómicas aplicando chi-cuadrado. En el ejemplo ( $\alpha = 0.05$ ,  $X^2$ :  $p = 0,004$ ), en sujetos valorados óptimamente, se observa una diferencia de proporciones de 0,39 (IC 95%: 0,16 - 0,62) entre tomar levotiroxina en ayunas o con alimentos.

Conclusión principal:

Es útil la distribución binomial en investigación en problemas de salud, en circunstancias particulares.

Palabras clave: Distribución binomial. Hipotiroidismo. Levotiroxina.

## CONTENIDO

Introducción

1. Justificación de la propuesta expuesta en este artículo.
2. Material y métodos.

- 2.1 Métodos del ejemplo didáctico propuesto.
- 2.2 Distribución binomial.
3. Resultados.
4. Discusión.
5. Conclusiones.
6. Referencias bibliográficas.

## 1. Justificación de la propuesta expuesta en este artículo.

En este artículo se presentará a la distribución binomial como herramienta útil en investigaciones donde se necesita establecer un criterio de grado de control sobre problemas de salud, difícil de obtener en la evidencia publicada y de formularse operativamente. Un sector investigador minoritario emplea la distribución binomial, pero se intentará su adopción por un amplio número de investigadores tras su lectura.

El artículo es teórico. En él, se explicará la distribución de probabilidad binomial y su correcta aplicación. Pero, dado el contenido matemático expuesto y como apoyo pedagógico para su mejor comprensión, se apoyará con un ejemplo dentro del contexto de una investigación en salud: en hipotiroidismo primario, evidenciar en una investigación de diseño longitudinal una variable influyente en la valoración óptima del tratamiento de sustitución con levotiroxina: la tendencia de ingestión de la dosis de levotiroxina, o bien en ayunas o bien junto con alimentos.

El hipotiroidismo primario aumenta la hormona tiroestimulante (TSH). En España, la prevalencia del hipotiroidismo alcanza un elevado porcentaje (10%) siendo el hipotiroidismo primario el 9% del anterior, siendo tratado con medicamentos el 4,2%. El tratamiento de sustitución oral con levotiroxina intenta restaurar el estado eutiroideo normalizando la concentración sérica de TSH. El control clínico del tratamiento de sustitución se efectúa evaluando puntualmente la concentración sérica de TSH. Tras la primera dosis o modificaciones posteriores, debe esperarse un periodo de 4-6 semanas para valorar su efecto. Concretada la dosis adecuada, los controles pueden espaciarse a medio o un año.<sup>1</sup>

Muchas variables afectan la estabilidad del control analítico de TSH, como se refleja en la tabla 1.

Variables que aumentan las necesidades de levotiroxina	
Incumplimiento terapéutico	
Fármacos que disminuyen la absorción de levotiroxina	Sulfato ferroso, resinas de intercambio iónico, Omeprazol, lansoprazol, orlistat, raloxifeno, colestiramina, colestipol, antiácidos con hidróxido de aluminio, manganeso
Fármacos que aumentan el metabolismo de la tiroxina	Rifampicina, carbamacepina, fenobarbital, simvastatina, lovastatina
Interacción con alimentos que originan la mala absorción de levotiroxina	Soja y derivados, dietas con alto porcentaje de fibra vegetal o ricas en grasas, café, nueces, ciruelas.
Enfermedades concomitantes	Celiaquía, gastritis atrófica, cirugía gastrointestinal (bypass yeyunoileal)
Otras	Dosis inadecuada, error en dispensación farmacéutica, gestación.

Tabla 1. Variables recogidas en la bibliografía que afectan la estabilidad del control analítico de TSH.

No obstante, la ingestión de la dosis diaria de levotiroxina en ayunas o simultáneamente con alimentos no se identifica como variable de influencia. La bibliografía: Agencia Española del Medicamento (2020), Vademecum Internacional (2008), Medimecum (2016), indica la ingestión en ayunas, mencionándose efectos negativos en asociación con alimentos.<sup>2,3,4</sup>

Como hipótesis de trabajo interesaría conocer si, en investigaciones de diseño longitudinal, la ingestión de la dosis diaria de levotiroxina en ayunas o simultáneamente con alimentos, pudiera identificarse como variable ventajosa para obtener una valoración óptima del control sérico de TSH. Los autores desarrollarían una investigación cuyo objetivo principal sería estudiar una asociación entre la ingesta de la dosis de levotiroxina en ayunas y obtener una valoración óptima del control analítico de TSH. El objetivo específico, por ejemplo, sería calcular en sujetos con prescripción de levotiroxina, con control analítico de TSH valorado como óptimo, una diferencia de proporciones  $\geq 20$  entre aquellos que ingieren la dosis en ayunas y los que la ingieren junto con alimentos, en un periodo de tiempo y zona geográfica concretos.

¿Cómo plantear este problema? Estimar en una población la calidad del control del tratamiento de sustitución podría hacerse fácilmente en un momento concreto planificando un diseño transversal y prueba de  $X^2$  (chi-cuadrado de independencia). Pero el resultado obtenido solamente valdría para ese momento y, al discutirlo, no podrían identificarse o descartarse otras influencias. Si la variable de interés de la investigación, “tendencia de ingestión de la dosis de levotiroxina, en ayunas o junto con alimentos”, influye en la valoración óptima del tratamiento (la cual se definirá posteriormente), dicha variable debería estudiarse a largo plazo y no transversalmente. Una investigación más ambiciosa, usando igualmente  $X^2$ , ampliaría el número de analíticas de control por sujeto realizadas en un período de tiempo

indefinido. Pero incluir un número impreciso de determinaciones de TSH, diferente para cada sujeto de una muestra y durante un periodo indeterminado complica el diseño: un sujeto puede tener 2 analíticas en 2 años, otro 4 en 1, otro 8 en 5... ¿Cómo equiparar la valoración de un sujeto con 4 analíticas con la de otro con 11, y considerando, además, diferentes periodos de tiempo? En el ejemplo de investigación propuesto se analizarían dos variables dicotómicas: “tendencia de ingestión de la dosis de levotiroxina (en ayunas o junto con alimentos)”, y “valoración del control analítico de TSH (óptima o subóptima)”. Operativizar la primera variable sería sencillo, preguntando al sujeto su tendencia habitual de ingestión de la dosis, en ayunas o no. Operativizar la segunda variable requiere especificar, para el fin de esta investigación, qué se entiende por un tratamiento de sustitución valorado como óptimo o subóptimo y, como definir operativamente dicha variable dicotómica. Esta última operativización sería factible mediante la distribución binomial.

El objetivo principal del artículo es fomentar en investigadores el conocimiento y uso de la distribución de probabilidad binomial en problemas de salud con dificultades similares al del ejemplo de investigación propuesto.

## 2. Material y métodos

En el ejemplo de investigación sugerido una posible estrategia para valorar óptimamente, en un sujeto, un tratamiento de sustitución, consistiría en cuantificar un porcentaje mínimo de controles analíticos con resultado de TSH normal con respecto al total de controles analíticos realizados durante el periodo de tiempo transcurrido desde el análisis con fecha más actual hasta la fecha de la primera analítica realizada después de la analítica de control de las 6 semanas (realizada ésta tras la prescripción de la primera o la última dosis de levotiroxina). A este periodo de tiempo lo llamaremos a partir de ahora “periodo de tiempo considerado”. Así, un sujeto tiene 3 analíticas con resultado normal de un total de 5 analíticas de control realizadas en 3 años, otro caso 4 de 8 analíticas totales en 4 años, u otro 11 de 12 en 7 años. El número total de controles analíticos realizados por un solo sujeto durante el periodo de tiempo considerado,  $N$ , sería una variable a tener en cuenta. Podría consensuarse que el porcentaje de analíticas con resultado de TSH normal con respecto al total,  $N$ , realizadas en el período de tiempo considerado sea  $\geq 90\%$ . Se detallará posteriormente como cuantificar en un solo sujeto dicho porcentaje. Escoger como criterio de valoración óptimo de tratamiento un punto de corte de porcentaje  $\geq 90\%$  será justificado en la discusión.

Primeramente, se describirá el método en el ejemplo de investigación propuesto y, posteriormente, se presentará la distribución de probabilidad binomial como utilidad

para operativizar justificadamente una variable dicotómica clave para el diseño de investigación.

## 2.1 Métodos del ejemplo didáctico propuesto

Diseño de investigación observacional, analítico, retrospectivo. La población de estudio serían sujetos con hipotiroidismo primario y tratamiento de sustitución con diferentes dosis prescritas de levotiroxina, con otras patologías y medicamentos asociados. Se seleccionaría una muestra de tamaño, considerando una potencia estadística del 80% en una prueba de chi-cuadrado de independencia con un nivel de significación  $\alpha = 0,05$ , calculada en 58 sujetos. Detectaría en los sujetos con valoración óptima del control analítico de TSH  $\geq 90\%$ , una diferencia de proporciones  $\geq 0,20$  entre aquellos que ingieren el fármaco en ayunas y los que lo ingieren con alimentos.

La obtención de los datos se formalizaría. Mediante un cuestionario estructurado y abierto se preguntaría a cada sujeto su tendencia de ingestión. Como convenio se consideraría la tendencia de ingestión de la dosis de levotiroxina en ayunas, a la conducta de no ingerir alimento sólido o líquido, excepto agua, durante un tiempo mínimo de 30 minutos después de tomar levotiroxina.

Posteriormente, se valoraría, en un sujeto, el control analítico de TSH (óptima o subóptima). Se contabilizaría el total de analíticas de control de TSH a partir del periodo de tiempo considerado. Cada control analítico de TSH podría tener dos resultados: normal (cuantificado dentro del intervalo normal) o anómalo (fuera del mismo). Se calcularía el porcentaje de controles analíticos de TSH con resultado normal con respecto al total, N, realizados durante el periodo de tiempo considerado. Se descartaría a sujetos con menos de 4 analíticas de control desde el periodo de tiempo considerado (motivo justificado en la discusión).

En general, para valorar un porcentaje  $\geq 90\%$  se cuantificarían las analíticas realizadas en el período de tiempo considerado, N, que supondrán el 100%. Una simple regla de tres proporcionaría el número mínimo de analíticas con resultado de TSH correcto para alcanzar un porcentaje  $\geq 90\%$ :

$$\begin{array}{l} 90 \rightarrow 100 \\ X \rightarrow N \end{array} \quad \text{por tanto, } X = (90 \cdot N) / 100 =$$

Por ejemplo, con 5 o 10 analíticas,

$$\begin{array}{l} 90 \rightarrow 100 \\ X \rightarrow 5 \end{array} \quad \text{así, } X = (90 \cdot 5) / 100 = 4,5 \text{ analíticas.}$$

El número mínimo de analíticas con resultado de TSH correcto que estima un porcentaje  $\geq 90\%$  será 5, un número entero.

$$\begin{array}{l} 90 \rightarrow 100 \\ X \rightarrow 10 \end{array} \quad \text{así, } X = (90 \cdot 10) / 100 = 9 \text{ analíticas.}$$

Aquí, el número de analíticas con resultado de TSH correcto que estima un porcentaje  $\geq 90\%$ , serán 9 o 10 analíticas.

La tabla 2 estima el número de analíticas con resultado de TSH correcto, r, sobre un total N realizados por un sujeto, necesarios para obtener un porcentaje  $\geq 90\%$ .

N	r
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	9 o 10
11	10 o 11
12	11 o 12
13	12 o 13
14	13 o 14
15	14 o 15
16	15 o 16

Tabla 2. Número de analíticas con resultado de TSH correcto, r, sobre un total N realizados por un sujeto, necesarios para obtener un porcentaje  $\geq 90\%$ .

Se establecería un criterio de valoración óptima: un porcentaje  $\geq 90\%$  de controles analíticos de TSH con resultado normal con respecto al total de controles analíticos realizados durante el periodo de tiempo considerado. En un sujeto con un porcentaje  $\geq 90\%$  del total, se valoraría el control analítico de TSH como óptimo. En caso contrario se valoraría como subóptimo. Con este procedimiento podría ser

operativa la variable dicotómica “valoración del control analítico de TSH de un sujeto (óptima o subóptima)”.

Las variables más importantes en la investigación propuesta se detallan en la tabla 3:

<p>Variable: Hora de toma de la levotiroxina.  Definición: hora en que el sujeto manifiesta que toma la dosis de levotiroxina.  Valores: Hora (horario de 24 horas) y minutos.  Variable de fecha.</p>
<p>Variable: Hora de ingestión de alimentos.  Definición: según la respuesta de hora de toma del fármaco, hora en que el sujeto manifiesta que ingiere alimento: comida, cena, etc.  Valores: hora (en horario de 24 horas) y minutos.  Variable de fecha.</p>
<p>Factor de estudio: Tendencia de ingestión de la dosis de levotiroxina.  Definición: nos indica si el paciente ingiere la dosis de levotiroxina en ayunas o simultáneamente con alimentos. Se basa en la información registrada en el cuestionario.  Categorías:  1: Ingestión en ayunas de la dosis de levotiroxina. Abstinencia de todo alimento, excepto agua, un tiempo mínimo de media hora después de tomar la dosis de levotiroxina.  2: Ingestión simultánea de la dosis de levotiroxina con alimentos. Ingestión de alimento en menos de media hora tras la toma de la dosis de levotiroxina.  Escala de medida: Escala nominal. Variable cualitativa dicotómica.</p>
<p>Variable: Valoración del valor de la TSH dentro o fuera del intervalo de referencia establecido por el hospital de referencia.  Definición: determina la categoría en la que se incluye a cada analítica en función de que su valor de la TSH esté dentro o fuera del intervalo de referencia establecido por el hospital de referencia (0,27 – 5 <math>\mu</math>U/ml).  Categorías:  1: Normal. Valor de la TSH dentro del intervalo de referencia establecido por el hospital de referencia.  0: Anormal. Valor de la TSH fuera del intervalo de referencia establecido por el hospital de referencia.  Escala de medida: Escala nominal. Variable cualitativa dicotómica.</p>
<p>Variable: Valoración del control analítico de la TSH de un caso válido.  Definición: determina la categoría en la que se incluye a cada sujeto en función del porcentaje de analíticas con resultado normal sobre el total de analíticas realizadas.  Categorías:  1: Óptimo (Porcentaje <math>\geq</math> 90 %). Valoración óptima.  2: Subóptimo (Porcentaje &lt; 90 %). Valoración subóptima.  Escala de medida: Escala nominal. Variable cualitativa dicotómica.</p>

Tabla 3. Descripción de las variables más importantes en la investigación propuesta como ejemplo pedagógico.

## 2.2 Distribución binomial.

Muchas situaciones de la vida real muestran un conjunto de características comunes. Determinando ese conjunto a una variedad de situaciones empíricas, puede formularse una distribución de probabilidad aplicable a esa variedad de situaciones. Estos modelos probabilísticos son funciones matemáticas conocidas de determinados tipos de experimentos aleatorios de variables aleatorias que describen algunos acontecimientos de la vida real y permiten calcular la probabilidad de que un experimento aleatorio concreto se produzca de una forma particular. Algunos corresponden a variables aleatorias discretas, como la distribución de probabilidad binomial.

Antes, se recordará qué es un experimento de Bernoulli: todo experimento aleatorio en el que solo pueden ocurrir dos sucesos mutuamente excluyentes: éxito o fracaso. Por ejemplo, lanzar una moneda resulta cara (éxito) o cruz (fracaso).

Dado un experimento en el que  $X_1$  es una variable aleatoria tipo Bernoulli que toma los valores 1 para el éxito y 0 para el fracaso. Sea  $X_2$  la misma variable aleatoria definida en una nueva representación del mismo experimento. De la misma forma definir las variables aleatorias  $X_3, X_4, \dots, X_n$ , en sucesivas repeticiones del mismo experimento. Entonces, se define la variable binomial “suma de éxitos que se obtienen en  $n$  repeticiones del experimento”:

$$X = X_1 + X_2 + X_3 + \dots + X_n$$

Esta variable binomial es la suma de éxitos de las  $N$  variables tipo Bernoulli, que valdrá: 0, cuando todas las  $X_i$  tomen el valor 0; 1, cuando una de ellas tome el valor 1 y las restantes el valor 0; 2, cuando dos tomen el valor 1 y las restantes 0, ..., hasta  $N$ , cuando todas tomen el valor 1. Después, puede calcularse cuál es la probabilidad asociada a cada uno de los valores de la variable  $X$ . La función  $f(x)$  que asigna probabilidades a la variable “número de éxitos en  $N$  ensayos tipo Bernoulli”, se denomina función de probabilidad binomial.

En la investigación propuesta, para un sujeto, un experimento sería observar una analítica de control de TSH en el que  $X_1$  es una variable aleatoria tipo Bernoulli que da como resultado el suceso 1 (resultado con TSH normal) o 0 (resultado anómalo).  $X_2$  es la misma variable aleatoria definida en una observación de una segunda analítica. Igualmente, repetimos las variables aleatorias  $X_3, X_4 \dots X_n$  en sucesivas observaciones del resto de analíticas realizadas por un sujeto durante el

período de tiempo considerado. Entonces se define la variable binomial como “suma de analíticas de TSH con resultado normal que se obtienen observando las N analíticas de control de TSH realizadas por un sujeto, durante el período de tiempo considerado”:

$$X= 1+1+0+0+1= 3$$

En este caso, la variable toma el valor de 3 analíticas de control de TSH con resultado normal con respecto al número total de analíticas, 5.

$$X= 1+1+1+1+1= 5$$

Ahora, la variable toma el valor de 5 analíticas de TSH con resultado normal con respecto al número total de analíticas, 5.

La distribución binomial es la distribución de probabilidad de N variables aleatorias independientes tipo Bernoulli en las que la probabilidad de éxito, p, se mantiene constante en cada ensayo. Un experimento binomial tiene tres condiciones:

- Consiste en N pruebas independientes (número de ensayos).
- Cada ensayo tiene dos posibles resultados: 1 (éxito) o 0 (fracaso).
- La probabilidad de éxito, p, permanece constante en cada ensayo.

Este modelo de distribución se simboliza por B(N,p) y soluciona el siguiente problema: si al realizar un experimento aleatorio hay una probabilidad P de que ocurra un determinado suceso (éxito) ¿cuál es la probabilidad de que en N ensayos independientes se confirme el mismo suceso x veces? Esta pregunta se responde conociendo su función de probabilidad.

Si representamos por X la variable aleatoria binomial que representa el número de éxitos obtenidos en las N pruebas del experimento, la función de probabilidad de la distribución binomial que asigna probabilidades a la variable número de éxitos en N ensayos es:

$$f(x) = P(\text{obtener } r \text{ éxitos}) = P(X = r) = \binom{N}{r} p^r \cdot q^{N-r}; \quad \text{donde } \binom{N}{r} \text{ es un}$$

número combinatorio.

Siendo, P la probabilidad de obtener r éxitos, r la suma de éxitos y N el número total de ensayos realizados. La probabilidad asociada al suceso 1 es p, y q la probabilidad del suceso 0, siendo  $q = 1 - p$ .

La distribución binomial depende únicamente de los valores que tomen  $N$  y  $p$ , así que existe una familia de distribuciones binomiales. Una para cada par de parámetros de la distribución,  $N$  y  $p$ . Los cálculos pueden ser laboriosos. Existen tablas que proporcionan la probabilidad de que la variable  $X$  tome distintos valores según los diferentes valores de  $N$  y  $p$ .<sup>5,6,7,8</sup>

Continuando con el ejemplo de un sujeto con 5 analíticas de control, independientes, con  $p$  constante ( $p = 0,7$ ) en cada ensayo, se define una variable aleatoria binomial,  $X$ , como “número de analíticas de control de TSH con resultado normal”. Para la valoración óptima del control analítico de TSH de un sujeto será necesario alcanzar un porcentaje de analíticas con resultado normal  $\geq 90\%$  del total de analíticas, lo cual solamente se consigue, en un sujeto con 5 analíticas, si las 5 analíticas dan resultado normal. Así, la probabilidad de obtener 5 analíticas con resultado de TSH normal:

$$f(x) = P(\text{obtener 5 éxitos}) = P(X = 5) = \binom{5}{5} 0,7^5 \cdot 0,5^0 = 0,168$$

Siendo  $r = 5$  la suma de análisis con resultado de TSH normal y  $N = 5$  el número total de análisis realizados por un sujeto en el período de tiempo considerado. Nótese que, en este ejemplo, la probabilidad de éxito,  $p$ , permanece constante en cada ensayo. De la ecuación de la función de probabilidad binomial, en nuestro ejemplo, solo disponemos de 2 datos constantes:  $p$  y  $q$ . La tabla 2 mostró el número mínimo de analíticas con resultado de TSH correcto,  $r$ , necesarias para obtener un porcentaje  $\geq 90\%$  del total  $N$ .

### 3. Resultados.

En la investigación propuesta, primeramente, se exploraría (tabla 4) las particularidades de la muestra desconociendo, a priori, el porcentaje de sujetos que tomen levotiroxina en ayunas.

Edad	Media: 64 años Desviación típica: 7 Mínimo: 31 Máximo: 92 Rango: 61
Sexo	Varón: 0,38. Mujer: 0,62.
Tendencia de ingestión de la dosis de levotiroxina entre los pacientes en el periodo de tiempo de recogida de datos del estudio (proporción).	En ayunas: 0,63. Acompañado de alimentos: 0,37.

Tabla 4. Exploración de las variables de la muestra seleccionada en el ejemplo pedagógico.

La tabla de contingencia que cruza las dos variables, “tendencia de ingestión de la dosis de levotiroxina, en ayunas o junto con alimentos” y “valoración del control analítico de TSH de un sujeto, óptima o subóptima” se muestra en la tabla 5.

		Tendencia de ingestión de la dosis de levotiroxina.	
		Ingestión en ayunas de la dosis de levotiroxina.	Ingestión simultanea de la dosis de levotiroxina con alimentos.
Valoración del control analítico de la TSH	Óptimo: Porcentaje $\geq 90\%$	Recuento: 25 Proporción de pacientes que ingieren la dosis de levotiroxina en ayunas: 0,67	Recuento: 6 Proporción de pacientes que ingieren la dosis de levotiroxina en ayunas: 0,28
	Subóptimo: porcentaje $\leq 90\%$	Recuento: 12 Proporción de pacientes que ingieren la dosis de levotiroxina junto con alimentos: 0,33	Recuento 15 Proporción de pacientes que ingieren la dosis de levotiroxina junto con alimentos: 0,72
Total: 58		Recuento: 37	Recuento: 21

Tabla 5. Tabla de contingencia que cruza dos variables: “tendencia de ingestión de la dosis de levotiroxina, en ayunas o junto con alimentos” y “valoración del control analítico de TSH de un sujeto, óptima o subóptima”.

El contraste de hipótesis mediante la prueba estadística chi-cuadrado de independencia para dos variables cualitativas rechazaría la hipótesis nula:

$$A = 0.05, X^2(1, n = 58) = 8,188, p = 0,004.$$

Y, en los sujetos con valoración óptima del control analítico de TSH se observaría una diferencia de proporciones, entre aquellos que toman el fármaco en ayunas y los que lo toman junto con alimentos, de 0,39 (IC 95%: 0,16 - 0,62).

#### 4. Discusión.

Encontrar un método para alcanzar el objetivo del ejemplo de investigación propuesto supone responder a 2 preguntas: ¿Cómo se puede definir un criterio objetivo de valoración del control analítico de TSH de un sujeto como óptimo? Y ¿cómo establecer un punto de corte no arbitrariamente escogido, sino muy justificado, para diferenciar un caso con valoración de control óptimo de otro caso con valoración no óptima?

Para responder a estas preguntas se establecería un porcentaje  $\geq 90\%$  de controles analíticos de TSH con resultado normal con respecto al total de controles analíticos realizados durante el periodo de tiempo considerado. Se calcularía como se describió en el apartado métodos. Ya definido, ¿por qué elegir un 90%, y no un 60% o un 70% u 80% como cifra de corte justificada? Para su justificación se recurriría al modelo de distribución de probabilidad binomial y definiendo la variable aleatoria binomial “número de análisis con resultado de TSH normal”. El modelo de distribución de probabilidad binomial  $B(N,p)$  dependería de 2 parámetros:

- N: número total, N, variable, de análisis realizados por un sujeto durante el período de tiempo considerado.
- p: 0,7. En el ejemplo propuesto, el cálculo práctico de la probabilidad, p, asociada al suceso 1 de cada experimento  $X_i$  sería igual al cociente entre el número de casos favorables de ocurrencia del suceso 1 (tomando solamente la última analítica realizada por cada sujeto, suma de aquellas con resultado normal de TSH) y el número de casos posibles (tomando solamente la última analítica realizada por cada sujeto, suma de todas las analíticas).

Como  $p = 0,7$  permanece constante, existe una familia de distribuciones en función de  $N$ . Así, para un sujeto, con un total de análisis,  $N$ , realizados, se puede calcular la probabilidad asociada para un valor  $r$  (suma de análisis con resultado de TSH normal) que supone obtener un porcentaje  $\geq 90\%$ :

N	r	p asociada
2	2	$P(X=2) = 0,490$
3	3	$P(X=3) = 0,343$
4	4	$P(X=4) = 0,240$
5	5	$P(X=5) = 0,168$
6	6	$P(X=6) = 0,118$
7	7	$P(X=7) = 0,082$
8	8	$P(X=8) = 0,058$
9	9	$P(X=9) = 0,040$
10	9 o 10	$P(X=9) + P(X=10) = 0,040 + 0,028 = 0,068$
11	10 o 11	$P(X=10) + P(X=11) = 0,028 + 0,020 = 0,048$
12	11 o 12	$P(X=11) + P(X=12) = 0,020 + 0,013 = 0,033$
13	12 o 13	$P(X=12) + P(X=13) = 0,013 + 0,009 = 0,022$
14	13 o 14	$P(X=13) + P(X=14) = 0,009 + 0,006 = 0,015$
15	14 o 15	$P(X=14) + P(X=15) = 0,006 + 0,004 = 0,010$
16	15 o 16	$P(X=15) + P(X=16) = 0,004 + 0,003 = 0,007$

Tabla 6. Para un sujeto concreto, se muestra la probabilidad asociada para un valor  $r$  (suma de análisis con resultado de TSH normal) que supone obtener un porcentaje  $\geq 90\%$ , con un total de análisis,  $N$ , realizados en el periodo de tiempo considerado.

Para un sujeto, y en función del número total de analíticas realizadas en el periodo de tiempo considerado, se comprueba que las probabilidades son suficientemente pequeñas como para descartar la posibilidad de que la valoración óptima del control analítico de TSH se deba al azar; y disminuyen conforme aumenta el número de análisis realizados. En el ejemplo propuesto, la media de analíticas realizadas por sujeto sería 6, siendo 4, 5, 6, 7 y 9 las categorías más frecuentes. Para esos valores de  $N$  (figura 6) las probabilidades serían pequeñas, se descartaría el azar en la obtención de la valoración óptima del control de TSH. Comparando, si en la operativización de la variable se estableciera como punto de corte un porcentaje  $\geq 70\%$ , la probabilidad asociada para un sujeto con 10 analíticas de obtener un porcentaje  $\geq 70\%$  sería de 0,64, probabilidad muy alta de ocurrir por azar. Estableciendo el porcentaje  $\geq 90\%$ , en el mismo sujeto la probabilidad sería de 0,06, probabilidad muy baja de obtener por azar. Concluyendo, este razonamiento (con ayuda de la distribución binomial) motivaría la justificación del porcentaje  $\geq 90\%$  como valor de corte en la operativización de la variable.

En la investigación propuesta, la probabilidad de que el suceso ocurra por azar es prácticamente despreciable, lo que induce a considerar variables cuya influencia facilite alcanzar dicho porcentaje. E identificándose una candidata: la ingestión de la dosis diaria de levotiroxina en ayunas, la cual no es distinguida en la bibliografía consultada como una variable influyente en la estabilidad de la valoración del control analítico de TSH. El diseño del estudio implicaría utilizar la prueba estadística  $X^2$  y el cruce de dos variables dicotómicas. Pero categorizar como dicotómica la variable “tendencia de ingestión de la dosis de levotiroxina”, sería sencillo procediendo tal y como se planteó en el apartado métodos. Categorizar la otra variable como dicotómica plantearía problemas de definición, operativización y cuantificación de la variable. No obstante, con la estrategia utilizada podría realizarse válida y justificadamente. Y, gracias a ello, consecuentemente, en el ejemplo se demostraría una asociación de las 2 variables tras rechazar la hipótesis nula:  $\alpha = 0.05$ ,  $X^2(1, n = 58) = 8,188$ ,  $p = 0,004$ . Y, en sujetos con valoración óptima del control analítico de TSH se observaría una diferencia de proporciones, entre aquellos que toman el fármaco en ayunas y aquellos que lo toman con alimentos, de 0,39 (IC 95%: 0,16 - 0,62). Gracias al planteamiento del diseño y con la utilización de la distribución binomial, se conseguirían alcanzar los objetivos general y específico del ejemplo de investigación propuesto.

Utilizando la estrategia sugerida, para un sujeto, independientemente de N, sería baja la probabilidad asociada de obtener por azar una valoración óptima del tratamiento, garantía adicional que fortalecería el estudio. La probabilidad asociada a 2 o 3 controles analíticos (descartado ya el de las 6 semanas) es de 0,5 y 0,34, respectivamente, probabilidades muy altas de obtener por azar un resultado valorado óptimamente. Por ello, se descartaría a los sujetos con menos de 4 analíticas tras la prescripción de la última dosis.

Una fuente bibliográfica consultada, observacional, descriptiva, transversal, descartó la forma de administración de levotiroxina como factor influyente en el control hipotiroideo, éste definido con fines clínicos. Su diseño, con una única analítica de control, impide identificar, a priori, a sujetos mal controlados en el momento de la extracción sanguínea.<sup>9</sup> Controlar esta adversidad exige varios análisis de seguimiento en el tiempo como se plantea en la investigación propuesta. Ésta, estudiaría la influencia de la forma de administración en la valoración óptima del tratamiento de sustitución sin una finalidad clínica, solo de investigación, siendo retrospectivo e independiente del número de analíticas y del periodo de tiempo de su extracción. A favor de la causalidad de la asociación entre ambas variables se razonaría que el antecedente de ingerir el fármaco en ayunas influiría en la valoración del tratamiento hipotiroideo, y no al contrario. Siendo la investigación observacional (no experimental), con endogamia geográfica de la muestra seleccionada, y una información obtenida con el cuestionario no confrontada (datos veraces o no), posibles variables de influencia no serían controladas. Sería un estudio original que asociaría en un diseño longitudinal la estabilidad de los controles analíticos de TSH

(dentro del intervalo normal) con la ingestión de la dosis en ayunas. No aspiraría a sentar un criterio clínico, solamente apoyaría a la finalidad de investigar la influencia de una variable, la ingestión en ayunas de la dosis de levotiroxina, en la calidad del tratamiento.

## 5. Conclusiones

En ciencias de la salud, las distribuciones de probabilidad para variables continuas más conocidas y utilizadas son la distribución de probabilidad normal y la distribución chi-cuadrado. Incluso otras, como las distribuciones t de Student o la F de Snedecor, son conocidas, si bien no tan empleadas como las anteriores. Pero existen otros modelos de distribución de probabilidad.

Este artículo presentó un modelo para variables aleatorias discretas: el modelo de distribución de probabilidad binomial. Se expuso su definición matemática, y su correcta aplicación. El ejemplo de investigación propuesto (con un contraste de hipótesis  $X^2$ ) requería el cruce estadístico de dos variables dicotómicas. Se utilizó la distribución binomial para justificar científicamente convertir una variable teórica, “valoración del control analítico de la TSH de un caso válido, óptimo o subóptimo”, en una variable operativa al valorarse óptimamente el control analítico de TSH si el porcentaje de analíticas de control de TSH con resultado normal fuera  $\geq 90\%$  del total de analíticas realizadas por el sujeto durante el período de tiempo considerado y, con la cual, clasificar a los sujetos en dos grupos. La justificación del valor de corte propuesto ( $\geq 90\%$ ) se sustentaría en la reducida probabilidad de obtener dicho porcentaje, calculada con la ayuda de la función de probabilidad binomial, debido exclusivamente al azar. Así, el ejemplo de investigación propuesto podría plantearse metodológica y estadísticamente y demostrar la asociación entre la ingestión en ayunas de la dosis de levotiroxina y la estabilidad en el tiempo del control óptimo del tratamiento de sustitución en pacientes con tratamiento de sustitución.

Intentar identificar o evidenciar posibles variables de influencia en la valoración del control en otros problemas de salud, además del ejemplo expuesto, que afectan a los ciudadanos (diabetes, hipertensión, dislipemias...), definidos como determinados tipos de experimentos aleatorios de variables aleatorias discretas y con posibilidad de cálculo de la probabilidad de que un experimento se produzca de una forma particular, será posible apoyándose en un planteamiento similar al ejemplo propuesto y con ayuda de la distribución binomial, facilitándose el desarrollo y planteamiento metodológico de diseños de investigación, que pudieran no ser abordables con las estrategias metodológicas y estadísticas más frecuentemente utilizadas por el personal sanitario.

## 6. Referencias bibliográficas

1. Arroyo, Jaume; Gil, Pilar; Llauradó, Gemma; Caixàs, Assumpta. Hipotiroidismo primario: consideraciones para una buena utilización del tratamiento con levotiroxina. *Medicina Clinica (Barc)*. 2011;136(5):207–214. Disponible en: <https://www.elsevier.es/es-revista-medicina-clinica-2-articulo-hipotiroidismo-primario-consideraciones-una-buena-S0025775309010070> [acceso:08/12/2021]
2. Agencia Española de Medicamentos y Productos Sanitarios. Ficha Técnica Eutirox® 2022. Disponible en: [https://cima.aemps.es/cima/pdfs/es/ft/64011/FichaTecnica\\_64011.html.pdf](https://cima.aemps.es/cima/pdfs/es/ft/64011/FichaTecnica_64011.html.pdf) [acceso:09/12/2021]
3. Villa, Luís. *Medimecum 2016 Guía de Terapia Farmacológica*. 2016. (21 ed.)p.533-534.
4. Canales, María Jesús. *Vademecum Internacional 2008*. CMP Medicom Ed.S.A. 2008 (49 ed.)p 1324-1325.
5. Quintanilla, Francisco. *Matemáticas Fundamentales*. Marcombo, S.A. 2014 (1ª ed.).
6. Bonnet Jerez, Jose Luis. *Lecciones de Estadística. Estadística Descriptiva y Probabilidad*. Editorial Club Universitario, 2003(1ª ed.)
7. Martínez Arias, María Rosario; Maciá Antón, María Araceli. Distribuciones de Probabilidad de Variables Aleatorias Discretas. En: Seoane, Julio (editor). Psicología Matemática I. Tomo II. UNED;1985 (1ª ed.)p. 5-28
8. Padilla, Miguel; Merino, José María; Rodríguez-Miñon, Pedro; Moreno, Enrique. *Psicología Matemática I*. UNED; 1996.(1ª ed.)p. 273-478.
9. Sánchez Castillo, J., & López Gutiérrez, J. J. (2015). Fallo terapéutico de levotiroxina en el manejo de hipotiroidismo en una institución de salud de Bogotá, D.C. Un estudio analítico de prevalencia. *Revista Colombiana de Ciencias Químico-Farmacéuticas*, 2015;44(1),16–33. Disponible en: <https://revistas.unal.edu.co/index.php/rccquifa/article/view/54234> [acceso:12/01/2022]