

TEMAS DE MATEMÁTICAS

(Oposiciones de Secundaria)

TEMA 61

DESIGUALDAD DE TCHEBYCHEV. COEFICIENTE DE VARIACION. VARIABLE NORMALIZADA. APLICACIÓN AL ANÁLISIS, INTERPRETACIÓN Y COMPARACIÓN DE DATOS ESTADÍSTICOS.

1. Variable Aleatoria.
 - 1.1. Concepto de Variable Aleatoria.
 - 1.2. Variable Aleatoria Discreta.
 - 1.3. Variable Aleatoria Continua.
 2. Esperanza Matemática.
 - 2.1. Momentos.
 - 2.2. Variable Normalizada
 - 2.3. Coeficiente de Variación.
 3. Desigualdades de Markov y Tchebychev.
 4. Teoremas de Bernouilli y Moivre.
 5. Aplicación al Análisis, Interpretación y Comparación de Datos Estadísticos.
 - 5.1. Ejemplos.
- Bibliografía Recomendada.

TEMA 61

DESIGUALDAD DE TCHEBYSHEV. COEFICIENTE DE VARIACION. VARIABLE NORMALIZADA. APLICACIÓN AL ANÁLISIS, INTERPRETACIÓN Y COMPARACIÓN DE DATOS ESTADÍSTICOS.

1. VARIABLES ALEATORIAS.

1.1. Concepto de Variable Aleatoria.

DEF Sea un experimento E y sea Ω el espacio muestral asociado con el experimento. Llamaremos Variable Aleatoria a una función X que asigna a cada elemento $\omega \in \Omega$ un número real $X(\omega)$.

DEF Diremos que una variable aleatoria es Discreta si sólo toma un número finito de valores o un número infinito pero numerable.

Si la variable aleatoria es discreta, enumeraremos los valores que toma como x_1, x_2, x_3, \dots

DEF Diremos que una variable aleatoria es Continua si toma un número infinito no numerable de valores

1.2. Variables Aleatorias Discretas.

DEF Definimos la función de Probabilidad de la variable aleatoria discreta X como aquella función que asigna a cada número real x_i la probabilidad de que la variable aleatoria tome ese valor, $X=x_i$. $f(x_i) = P(X=x_i)$

Tengamos en cuenta que la variable aleatoria X estaba definida sobre un espacio muestral Ω , el cual es discreto. En cambio, la función de probabilidad $f(x)$ esta definida sobre \mathbb{R} .

Así pues, si el número real x no pertenece al conjunto de valores que toma X , se verificará

$$f(x_i) = P(X=x_i) = 0$$

En cambio, si pertenece al conjunto de los valores asumibles por X , se verificará:

$$f(x_i) = P(X=x_i) > 0$$

En este caso, $f(x_i)$ será la suma de las probabilidades de todos los sucesos elementales a los que X asigna el valor x_i .

Las propiedades fundamentales de la función de probabilidad no son más que la traducción de los axiomas de la probabilidad.

PROP Si x_1, x_2, \dots, x_n son los valores que toma la variable aleatoria X , se verifica:

- 1) $\sum f(x_i) = P(X = x_1) + P(X = x_2) + \dots + P(X = x_n)$
- 2) $f(x_i) > 0 \quad \forall i: 1, 2, \dots, n$
- 3) Si $a < b < c$, y definimos los sucesos mutuamente excluyentes $A = \{a \leq X \leq b\}$ y $B = \{b < X \leq c\}$. Si $C = \{a \leq X \leq c\}$, es claro que $C = A \cup B$ y, por tanto:

$$P(a \leq X \leq c) = P(a \leq X \leq b) + P(b < X \leq c)$$

DEF Definimos la función de Distribución de una variable aleatoria discreta X como $F(x) = P(X \leq x_i)$.

Esta claro que esa probabilidad será la suma de las probabilidades de todos los sucesos elementales de Ω a los que X asigna valores menores o iguales que x_i .

PROP La función de distribución de una variable aleatoria verifica las siguientes propiedades:

- 1) $F(-\infty) = 0$
- 2) $F(+\infty) = 1$
- 3) $F(x)$ es creciente.
- 4) $P(a < X \leq b) = F(b) - F(a)$.
- 5) $F(x)$ es Continua por la Derecha.

Dem.

- 1) Sean x_1, x_2, \dots, x_n los únicos valores que puede tomar la variable aleatoria X y sea x_0 un valor menor que el más pequeño de los n valores anteriores. Es claro que el suceso $\{X \leq x_0\}$ es un suceso imposible. Por tanto su probabilidad es cero, manteniendose dicho valor cuando más pequeño sea x_0 . Luego se verifica.

- 2) Análoga a la anterior.

- 3) Por propia Definición, ya que todos los sumandos son no negativos.

- 4) Los sucesos $\{X \leq a\}$ y $\{a < X \leq b\}$ son mutuamente excluyentes, y su unión es el suceso $\{X \leq b\}$

$$P(X \leq b) = P(X \leq a) + P(a < X \leq b)$$

Es decir

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$$

- 5) Por Definición.

1.3. Variables Aleatorias Continuas.

DEF Llamamos función de Densidad de Probabilidad de una variable aleatoria continua X a una función $f(x)$ que verifica las dos condiciones siguientes:

- 1) $f(x) \geq 0$
- 2) $\int_{-\infty}^{+\infty} f(x) dx = 1$

DEF Llamaremos función de Distribución de la variable aleatoria continua X a una función que asigna a todo número real x, la probabilidad de que X sea igual o menor que x.

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t)dt$$

PROP La función de distribución de una variable aleatoria continua X verifica:

- 1) $F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0$
- 2) $F(+\infty) = \lim_{x \rightarrow +\infty} F(x) = 1$
- 3) F(x) es una función no decreciente.
- 4) $P(a < X \leq b) = F(b) - F(a)$.
- 5) Si F(x) es derivable, $f(x) = \frac{dF(x)}{dx}$

Dem.

Inmediata.

2. ESPERANZA MATEMÁTICA.

DEF Sea X una variable aleatoria con función de probabilidad f(x). Llamaremos Esperanza Matemática de X a μ , siendo:

- 1) Caso Discreto: $\mu = E(X) = \sum_i x_i f(x_i)$
- 2) Caso Continuo: $\mu = E(X) = \int_{-\infty}^{+\infty} xf(x)dx$

La esperanza matemática, denotada por E(X), también recibe el nombre de Media o Valor Esperado.

Cuando una variable aleatoria se expresa mediante una función $Y=G(X)$, con X otra variable aleatoria, podemos expresar la esperanza matemática de Y utilizando X como sigue:

- 1) Caso Discreto. $\mu = E(Y) = E(G(X)) = \sum_i g(x_i) f(x_i) = \sum_i y_i h(y_i)$
- 2) Caso Continuo. $\mu = E(Y) = E(G(X)) = \int_{-\infty}^{+\infty} g(x) f(x)dx = \int_{-\infty}^{+\infty} yh(y)dy$

Si tenemos que una variable aleatoria Z se expresa como producto de otras dos variables aleatorias, $Z=X \cdot Y$, su esperanza matemática será:

$$E(Z) = E(XY) = \sum_{i,j} x_i y_j f(x_i, y_j)$$

$$E(Z) = E(XY) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xyf(x, y)dxdy$$

PROP Sean a y b constantes y X una variable aleatoria con media μ . Si $Y=aX+b$ entonces $E(Y) = a\mu+b$.

PROP El valor esperado de la suma o diferencia de dos o más funciones de una variable aleatoria X , es la suma o diferencia de los valores esperados de las funciones.

$$E[g(X)\pm h(X)] = E[g(X)] \pm E[h(X)]$$

PROP La Esperanza Matemática es una función lineal. $E(aX+bY) = aE(X) + bE(Y)$.

2.1. Momentos.

DEF Sea X una variable aleatoria Discreta con función de probabilidad $f(x)$. Llamaremos Momento de Orden r Respecto al Origen de la variable aleatoria X , y lo denotamos por α_r , a la expresión:

$$\alpha_r = E(X^r) = \sum x^r f(x)$$

DEF Sea X una variable aleatoria Continua con función de densidad de probabilidad $f(x)$. Llamaremos Momento de Orden r Respecto al Origen de la variable aleatoria X , y lo denotamos por α_r , a la expresión:

$$\alpha_r = E(X^r) = \int_{-\infty}^{+\infty} x^r f(x) dx$$

Podemos destacar a μ_1 que corresponde precisamente con $E(X)$, la media de la variable aleatoria.

DEF Sea X una variable aleatoria Discreta con función de probabilidad $f(x)$. Llamaremos Momento Central de Orden r de la variable aleatoria X , y lo denotamos por μ_r , a la expresión:

$$\mu_r = E[(X - E(X))^r] = \sum (X - E(X))^r f(x)$$

DEF Sea X una variable aleatoria Continua con función de densidad de probabilidad $f(x)$. Llamaremos Momento Central de Orden r de la variable aleatoria X , y lo denotamos por μ_r , a la expresión:

$$\mu_r = E[(X - E(X))^r] = \int_{-\infty}^{+\infty} (X - E(X))^r f(x) dx$$

DEF Sea X una variable aleatoria con distribución de probabilidades $f(x)$ y media μ . Definimos la Varianza de X , y se denota por $\text{Var}(X)$, como μ_2 , Momento Central de Orden 2.

En el caso discreto $\text{Var}(X) = \mu_2 = E[(X - E(X))^2] = \sum (x_i - \mu)^2 f(x_i)$

En el caso continuo $\text{Var}(X) = \mu_2 = E[(X - E(X))^2] = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx$

DEF Llamamos Desviación Típica o Estándar de una variable aleatoria X , y se denota por σ , a la raíz cuadrada de la Varianza.

$$s = \sqrt{\text{Var}(X)} = \sqrt{m_2}$$

PROP La varianza de una variable aleatoria X se puede expresar como la media de los cuadrados menos el cuadrado de la media.

$$\text{Var}(X) = E[(X - E(X))^2] = E(X^2) - E(X)^2$$

PROP Si a y b son constantes y X es una variable aleatoria con media μ y varianza σ^2 . Se verifica:

$$\text{Var}(aX+b) = a^2 \text{Var}(X)$$

2.2. Variable Normalizada.

DEF Si en la proposición anterior, tomamos como caso particular los valores $a = \frac{1}{s}$ y $b = -\frac{m}{s}$, la variable $\frac{X - m}{s}$, que expresa la desviación de la variable aleatoria X respecto de su media y medida en unidades de la desviación típica, recibe el nombre de Variable Normalizada o Tipificada.

Nótese que la media de la variable normalizada es cero y su desviación típica uno.

$$E\left(\frac{X - m}{s}\right) = 0 \quad \text{Var}\left[\left(\frac{X - m}{s}\right)\right]^{\frac{1}{2}} = \sqrt{E\left[\left(\frac{X - m}{s}\right)^2\right]} = 1$$

2.3. Coeficiente de Variación.

DEF Llamaremos Coeficiente de Variación de la variable aleatoria X al cociente de la desviación típica por la media.

$$\frac{s}{E(X)}$$

Para poder comparar las medias aritméticas de dos distribuciones que vengan dadas en unidades diferentes tenemos el coeficiente de variación de Pearson.

DEF Definimos el coeficiente de variación de Pearson como la relación por cociente entre la desviación típica y la media aritmética.

$$V = \frac{s}{m}$$

En primer lugar, tenemos que dicha medida es adimensional. En segundo lugar, V representa el número de veces que σ contiene a μ . Cuanto mayor sea V , más veces contendrá σ a μ , luego relativamente a mayor V menor representatividad de μ .

Este coeficiente se suele expresar en tanto por ciento, siendo

$$V = \frac{\mathbf{s}}{\mathbf{m}} \cdot 100$$

Como tanto en σ como en μ han intervenido todos los valores de la distribución, V presenta la garantía de que utiliza toda la información.

La cota inferior de V es cero, al ser éste el menor valor que puede tomar σ , y es el valor de V que indica la máxima representatividad de μ .

En caso de que la media aritmética sea nula, el valor de V no es significativo, ya que su resultado numérico nos puede hacer tomar conclusiones estadísticamente equivocadas.

3. DESIGUALDADES DE MARKOV Y TCHEBYCHEV.

En este apartado vamos a ver las desigualdades de Markov y Tchebychev. Ambas se basan en el concepto de valor esperado para establecer acotaciones sobre la probabilidad. La desigualdad de Markov establece una acotación de la probabilidad de una función no negativa de una variable aleatoria X . Con la desigualdad de Tchebychev, si conocemos la distribución de probabilidades de una variable aleatoria X , podemos calcular su esperanza $E(X)$ y su varianza $\text{Var}(X)$, si existen. Sin embargo, el recíproco es falso. Es decir, conociendo la media y la varianza de la variable aleatoria X no podemos reconstruir la distribución de probabilidades de X . Debido a esto, es conveniente obtener unas cotas superior e inferior para la función de probabilidad.

DESIGUALDAD DE MARKOV.

Dadas una función g no negativa de la variable aleatoria X y una constante t positiva, se verifica que

$$P[g(X) \geq t] \leq \frac{E[g(X)]}{t}$$

Dem.

La demostración vamos a realizarla para el caso de que la variable aleatoria X sea continua. El caso discreto es análogo.

Sea D el dominio en el que $g(X) \geq t$. Entonces:

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x)f(x)dx \geq \int_D g(x)f(x)dx \geq \int_D tf(x)dx \geq t \int_D f(x)dx$$

Teniendo ahora en cuenta las propiedades que verifica la función de densidad:

$$\int_{-\infty}^{+\infty} f(x)dx = 1 \quad \text{y} \quad \int_D f(x)dx = P[g(X) \geq t]$$

resulta que

$$E[g(X)] \geq tP[g(X) \geq t]$$

y operando llegamos a

$$P[g(X) \geq t] \leq \frac{E[g(X)]}{t}$$

Un caso particular que merece la pena destacar es cuando la función g sea la identidad, $g(X)=X$. En este caso

$$P[X \geq t] \leq \frac{E[X]}{t}$$

siempre que $X \geq 0$. Esta expresión nos va a permitir demostrar la desigualdad de Tchebychev.

DESIGUALDAD DE TCHEVYCHEV.

Sea X una variable aleatoria con media μ y varianza σ^2 . Entonces, $\forall k > 0$ se verifica:

$$P(\mu - k\sigma < X < \mu + k\sigma) = P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}$$

Dem.

Definamos la variable aleatoria $Y = [X - E(X)]^2$.

Como $P(Y \geq 0) = 1$ podemos aplicar la desigualdad de Markov, para obtener:

$$P(|X - \mu| \geq k\sigma) = P(Y \geq k^2 \sigma^2) \leq \frac{E(Y)}{k^2 \sigma^2} = \frac{\sigma^2}{k^2 \sigma^2} = \frac{1}{k^2}$$

verificando su complementario la desigualdad que queremos demostrar:

$$P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}$$

Veamos ahora otra manera de enunciar la desigualdad de Tchevybech, y vamos a realizar su demostración para el caso de una variable discreta.

DESIGUALDAD DE TCHEVYBECH.

Sea X una variable aleatoria. $\forall k > 0$ se verifica:

$$P(|X| \geq k) \leq \frac{E(X^2)}{k^2}$$

Dem.

Sea X una variable aleatoria discreta.

$$E(X^2) = \sum_i x_i^2 p_i = \sum_{x_i \leq -k} x_i^2 p_i + \sum_{-k < x_i < k} x_i^2 p_i + \sum_{k \leq x_i} x_i^2 p_i$$

Entonces

$$E(X^2) \geq \sum_{x_i \leq -k} x_i^2 p_i + \sum_{k \leq x_i} x_i^2 p_i \geq k^2 \left(\sum_{x_i \leq -k} p_i + \sum_{k \leq x_i} p_i \right)$$

$$E(X^2) \geq k^2 \cdot P(|X| \geq k) \Rightarrow P(|X| \geq k) \leq \frac{E(X^2)}{k^2}$$

Si en la desigualdad anterior sustituimos la variable aleatoria X por otra variable que esté tipificada $\frac{X - m}{s}$, obtenemos la desigualdad de Tchevychev demostrada en primer lugar.

La desigualdad de Tchevychev nos va a servir para poder justificar la introducción del concepto frecuencialista de probabilidad.

Sea A un suceso aleatorio, con probabilidad p , $P(A)=p$. Sea f la variable aleatoria que mide el número de apariciones de A en una serie de n observaciones independientes.

$$f = X_1 + X_2 + \dots + X_n \text{ donde } X_i = \begin{cases} 1 & \text{si se verifica } A \\ 0 & \text{si no se verifica } A \end{cases} \text{ en la } i\text{-ésima experiencia}$$

$$E(X_i) = 1 \cdot P(A) + 0 \cdot (1 - P(A)) = p$$

$$\text{Var}(X_i) = E[(X_i - p)^2] = p \cdot (1 - p)^2 + (1 - p) \cdot (0 - p)^2 = p \cdot (1 - p)$$

Como los experimentos son independientes:

$$E(f) = E(X_1) + E(X_2) + \dots + E(X_n) = np$$

$$\text{Var}(f) = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_n) = np(1 - p)$$

4. TEOREMAS DE BERNOULLI Y MOIVRE.

TEOREMA DE BERNOULLI.

Dado $k > 0$, la probabilidad de que la desviación absoluta de la frecuencia relativa de A sea mayor que k en la repetición de n experiencias independientes respecto a la probabilidad de A tiende a 0 cuando n tiende a ∞ .

Dem.

Apliquemos la desigualdad de Tchebychev a la variable aleatoria $\frac{f}{n}$.

$$E\left(\frac{f}{n}\right) = \frac{np}{n} = p \quad \text{y} \quad \sigma = \sqrt{\text{Var}\left(\frac{f}{n}\right)} = \frac{\sqrt{np(1-p)}}{n} = \sqrt{\frac{p(1-p)}{n}}$$

$$P\left(\left|\frac{f}{n} - p\right| \geq k\right) \leq \frac{p(1-p)}{nk^2} \leq \frac{1}{4nk^2} \quad \text{ya que} \quad p(1-p) \leq \frac{1}{4} \quad \forall p \in [0,1]$$

Entonces:

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{f}{n} - p\right| \geq k\right) = 0$$

Un caso algo más general fue obtenido por Poisson, que hacía variar la probabilidad de A de un experimento a otro.

Si la probabilidad de A en el experimento i-ésimo es p_i , tenemos que

$$E(f) = \sum_{i=1}^n p_i \quad \text{y} \quad \text{Var}(f) = \sum_{i=1}^n p_i(1-p_i)$$

y entonces

$$P\left(\left|\frac{f}{n} - \bar{p}\right| \geq k\right) \leq \frac{\sum_{i=1}^n p_i(1-p_i)}{n^2 k^2} \leq \frac{1}{4n^2 k^2}$$

siendo

$$\bar{p} = \frac{1}{n} \sum_{i=1}^n p_i$$

la probabilidad media.

El teorema de Bernouilli se publicó en el año 1713, 150 años después que la desigualdad de Tchebychev. Originalmente se demostró calculando directamente la probabilidad

$$P\left(\left|\frac{f}{n} - p\right| \leq k\right) = P(|f - np| \leq nk) = \sum_{n(p-k) \leq i \leq n(p+k)} \binom{n}{i} p^i (1-p)^{n-i}$$

Bernouilli tardó veinte años en calcular el sumatorio anterior, y probar que tiende a 1 cuando n tiende a ∞ .

Sin embargo, hoy en día puede deducirse un resultado mucho más general que el de Bernouilli y que fue dado por De Moivre en 1733, y que se podría enunciar como:

“La Distribución Binomial es Asintóticamente normal”

también conocido como Teorema Central del Límite, en forma reducida.

En la demostración de De Moivre, se utilizaba la variable tipificada

$$\mathbf{I} = \frac{f - np}{\sqrt{np(1-p)}}$$

Entonces:

$$P(\mathbf{I}_1 < \mathbf{I} < \mathbf{I}_2) = \sum_{np + \mathbf{I}_1 \sqrt{np(1-p)} \leq i \leq np + \mathbf{I}_2 \sqrt{np(1-p)}} \binom{n}{i} p^i (1-p)^{n-i}$$

Pues bien

$$\lim_{n \rightarrow \infty} P(\mathbf{I}_1 < \mathbf{I} < \mathbf{I}_2) = \frac{1}{\sqrt{2p}} \int_{\mathbf{I}_1}^{\mathbf{I}_2} e^{-\frac{u^2}{2}} du$$

siendo la expresión de la derecha la función de distribución normal.

El cálculo de la probabilidad

$$\sum_{np + \mathbf{I}_1 \sqrt{np(1-p)} \leq i \leq np + \mathbf{I}_2 \sqrt{np(1-p)}} \binom{n}{i} p^i (1-p)^{n-i}$$

anterior, cuando $n \rightarrow \infty$ puede hacerse de forma ligeramente distinta, lo que proporciona otra distribución límite, llamada de Poisson, muy útil cuando n es muy grande y p es muy pequeña.

Sea λ un parámetro tal que $\mathbf{I} = p \cdot n$ y $q = 1 - p = 1 - \frac{\mathbf{I}}{n}$

$$\begin{aligned} P(X = r) &= \binom{n}{r} \left(\frac{\mathbf{I}}{n}\right)^r \left(1 - \frac{\mathbf{I}}{n}\right)^{n-r} = \frac{n(n-1)\dots(n-r+1)}{n^r} \cdot \frac{\mathbf{I}^r}{r!} \left(1 - \frac{\mathbf{I}}{n}\right)^{n-r} = \\ &= \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{r-1}{n}\right) \cdot \frac{\mathbf{I}^r}{r!} \left(1 - \frac{\mathbf{I}}{n}\right)^n \cdot \left(1 - \frac{\mathbf{I}}{n}\right)^{-r} \end{aligned}$$

Tomando límites cuando $n \rightarrow \infty$ nos queda la distribución de Poisson

$$P(X = r) \approx \frac{\mathbf{I}^r}{r!} e^{-\mathbf{I}}$$

Se verifica que $\sum_{r=0}^{\infty} \frac{\mathbf{I}^r}{r!} e^{-\mathbf{I}} = 1$, la esperanza es $\lambda = np$ y la varianza es $\sigma^2 = \lambda$

En la práctica, se suele utilizar cuando es suceso A es raro y $np \leq 5$

5. APLICACIÓN AL ANÁLISIS, INTERPRETACIÓN Y COMPARACIÓN DE DATOS ESTADÍSTICOS.

La noción de ley de probabilidad correspondiente a una variable aleatoria se introduce en Estadística como un modelo de las regularidades que se observan al considerar series estadísticas.

Al igual que sucede con todas las teorías de las Matemáticas, la segunda parte es ver como se adaptan los modelos matemáticos a la realidad.

Igualmente, si llegamos a establecer criterios que permitan afirmar que una serie estadística se puede considerar como una cierta aproximación de una ley de probabilidad, podemos admitir que el mecanismo que conduce a estas observaciones será análogo al de los experimentos imaginados para obtener valores del universo que tiene dicha ley de probabilidad.

Concretamente, al aplicar métodos estadísticos para obtener nuevos conocimientos de los fenómenos naturales, se pueden considerar cuatro etapas:

a) Descripción.

Propone la recogida, clasificación y presentación resumida de datos relativos a un fenómeno.

b) Modelos.

Para explicar los hechos observados se formulan hipótesis, teorías o se buscan modelos, que expresen en forma matemática las relaciones que se han observado en los datos estadísticos.

c) Verificación.

También llamado Contraste del Modelo, Se realiza mediante la recogida de nuevos datos estadísticos relativos al fenómeno estudiado. Si la ley se confirma, podrá utilizarse en lo sucesivo. Si no, se descarta.

d) Predicción.

La teoría o modelo establecido permite establecer predicciones.

5.1. Ejemplos.

a) Ley de Mendel.

1) Descripción.

Mendel estudió el cruce de una variedad de guisantes amarillos con otros verdes. Los guisantes verdes, al reproducirse, dan siempre verdes, pero los amarillos dan unos sólo amarillos y otros amarillos y verdes. Estos amarillos dan una raza pura que da

indefinidamente amarillos. Si se cruzan verdes con amarillos de raza pura, se obtiene una primera raza de híbridos verdes. Si éstos se cruzan entre sí, se obtienen guisantes amarillos y verdes aproximadamente en una proporción 3 a 1.

2) Modelo Matemático.

Este modelo fue sugerido por el propio Mendel. En los cromosomas del guisante hay un corpúsculo portador del color. En la raza híbrida, unos gametos llevan el gen V y otros el A en la misma proporción. Al formarse las células, se pueden tener los siguientes tipos:

$$V_1V_2 \quad A_1V_2 \quad V_1A_2 \quad A_1A_2$$

Como A es dominante, entonces la proporción de amarillos es de $\frac{3}{4}$

3) Verificación.

En este caso, la comprobación de la Ley debe hacerse con test de hipótesis, por ejemplo la χ^2 de Pearson.

4) Predicción.

Una vez confirmada la ley, se puede saber con cierta probabilidad cual será el resultado del cruzamiento de dos plantas de guisantes en las condiciones anteriores.

b) Calidad en la Producción de Inyectables.

1) Descripción.

Se ha observado que una máquina produce inyectables con un porcentaje de defectuosos del 1% en un lote de 10.000 unidades.

2) Modelo.

El número de inyectables defectuosos en una caja de 200 unidades es variable, pero una teoría basada en la observación y en un modelo del cálculo de probabilidades permite considerar el número de defectuosos como una variable de Poisson

3) Verificación.

Si este modelo es confirmado por la experiencia, se puede utilizar para la predicción. Si no es confirmado, se revisa la hipótesis del paso 2.

4) Predicción.

La hipótesis y teoría anterior permite predecir que es prácticamente seguro que en una caja de 200 unidades de inyectables aparezcan, a lo sumo, cuatro defectuosos.

BIBLIOGRAFÍA RECOMENDADA.

Introducción a la Teoría de la Estadística. Aut.: Mood/Graybill. Ed. Aguilar.

Introducción a la Probabilidad y la Medida. Aut. Procopio Zoroa. Ed. PPU

Algoritmo. Matemáticas II. Cou. Aut.: Vizmanos y Anzola. Edit. SM.

Curso Básico de Estadística Económica. Aut.: M.P. Martín y F.J. Martín. Edit.: AC.