PREGUNTAS TEMA 1

1. *Definición de los momentos centrados y no centrados.*

Los momentos son valores calculados a partir de la distribución de frecuencias. Resultan muy útiles porque miden propiedades de la variable observada.

Momentos no centrados

Se define el momento no centrado (o respecto al origen) de orden *r* como:

$$a_r = \frac{1}{n}\sum_{i=1}^{k} x_i^r n_i = \sum_{i=1}^{k} x_i^r f_i \qquad \text{(para tablas con frecuencias)}$$

$$a_r = \frac{1}{n}\sum_{i=1}^{n} x_i^r \qquad \text{(para tablas sin frecuencias)}$$

- $a_0 = 1$ en cualquier distribución de frecuencias.
- El momento no centrado a_1 se conoce también como media aritmética y se nota como $\bar{x}$

Momentos centrados

Se define el momento centrado (o respecto a la media) de orden *r* como:

$$m_r = \frac{1}{n}\sum_{i=1}^{k} \left(x_i - \bar{x}\right)^r n_i = \sum_{i=1}^{k} \left(x_i - \bar{x}\right)^r f_i \qquad \text{(para tablas con frecuencias)}$$

$$m_r = \frac{1}{n}\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^r \qquad \text{(para tablas sin frecuencias)}$$

- $m_0 = 1$ y $m_1 = 0$ en cualquier distribución de frecuencias.
- El momento centrado m_2 se conoce también como varianza y se nota S^2
- $m_2 = a_2 - a_1^2$
- Análogamente, desarrollando el correspondiente binomio elevado a r, cualquier momento centrado puede escribirse en función de los momentos no centrados. Señalamos por su importancia los casos m_3 y m_4

$$m_3 = a_3 - 3a_2a_1 + 2a_1^3 \qquad\qquad m_4 = a_4 - 4a_3a_1 + 6a_2a_1^2 - 3a_1^4$$

2. *Media aritmética: Definición y propiedades.*

Es el valor característico de una serie de datos cuantitativos, este es el promedio más familiar y utilizado en los más diversos ámbitos. Trata de representar mediante un solo valor a un conjunto de datos y suele tomar una posición central respecto de los mismos, motivo por el que es conocida medidas de posición central.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i n_i = \sum_{i=1}^{k} x_i f_i \qquad \text{(para tablas con frecuencias)}$$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \text{(para tablas sin frecuencias)}$$

- Coincide con el momento no centrado $a_1 = x$.
- Consideramos n observaciones agrupadas en s conjuntos de datos con n_1, n_2, …,n_s observaciones cada uno y con medias x1 , x2, ..., xs respectivamente, entonces la media x de las n observaciones es:

$$\bar{x} = \frac{\bar{x}_1 n_1 + \ldots + \bar{x}_s n_s}{n_1 + \ldots + n_s} = \frac{1}{n}\sum_{i=1}^{s} \bar{x}_i n_i$$

- Con frecuencia se dividen o multiplican los valores de la variable por una constante, ex_i (cambio de escala) . En otras ocasiones se suma o resta una constante a los valores de la variable, $x_i + c$ (cambio de origen). Si realizamos una o ambas transformaciones crean una variable cuya media está relacionada con la media de la variable de partida según:

$$\bar{y} = e\bar{x} + c$$

3. *Diferencia entre medidas de dispersión absoluta y relativa. Defina las más importantes (varianza, coeficiente de variación,….).*

Las medidas de dispersión cuantifican la variabilidad o esparcimiento de los datos. Cuando esta dispersión se mide respecto de alguna medida de posición central (por ejemplo, la media) nos indica la mayor o menor representatividad de dicha medida

Los **recorridos, la varianza y la desviación típica** son medidas de **dispersión absoluta** que dependen de la unidad de medida de la variable (y por tanto les afecta el cambio de escala), lo anterior conlleva que no se puedan comparar estas medidas en dos variables con distinta unidad de medida.

Se prefiere trabajar con otras medidas de dispersión, obtenidas de las anteriores, que se denominan ***medidas de dispersión relativa***. Las medidas de dispersión relativa son adimensionales (no dependen de la unidad de medida de la variable estadística y por tanto no les afectan los cambios de escala)

Varianza:

Mide la dispersión o distancia de los datos, x_i , respecto de la media aritmética, $\bar{x}$. Esta medida está expresada en las unidades de los datos al cuadrado (p.e., €2, hab.2,...) por lo que no tiene una interpretación fácil, esta se define la desviación típica como la raíz cuadrada positiva de la varianza.

$$S^2 = \frac{1}{n}\sum_{i=1}^{k}\left(x_i - \bar{x}\right)^2 n_i \qquad \text{(para tablas con frecuencias)}$$

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 \qquad \text{(para tablas sin frecuencias)}$$

Coeficiente de variación:

Es la medida de dispersión relativa más utilizada. Se define como el cociente de la desviación típica sobre la media aritmética:

$$CV = \frac{S}{\bar{x}}$$

4. *Definición de varianza y propiedades.*

La varianza es una medida de dispersión que representa la variabilidad de una serie de datos respecto a su media. Formalmente se calcula como la suma de residuos al cuadrado divididos entre el total de observaciones. También se puede calcular como la desviación típica al cuadrado. Además la varianza coincide con el momento centrado de orden 2, $S^2 = m_2$.

La varianza mide la dispersión o distancia de los datos respecto de la media aritmética. Esta medida está expresada en las unidades de los datos al cuadrado por lo que no tiene una interpretación fácil. Con el objeto de tener una medida de dispersión expresada en las mismas unidades que los datos en estudio, se define la ***desviación típica*** como la raíz cuadrada positiva de la varianza.

En cuanto a las fórmulas de la varianza encontramos las siguientes:

$$S^2 = \frac{1}{n}\sum_{i=1}^{k}\left(x_i - \bar{x}\right)^2 n_i \qquad \text{(para tablas con frecuencias)}$$

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 \qquad \text{(para tablas sin frecuencias)}$$

Con frecuencia se dividen o multiplican los valores de la variable por una constante, ex_i, por ejemplo cuando decidimos expresar los valores en millones en lugar de en euros (en $ en lugar de en €,…). En otras ocasiones se suma o resta una constante a los valores de la variable, $x_i + c$. Si realizamos una o ambas transformaciones sobre la variable original obtenemos una nueva variable, $y_i = ex_i + c$, cuya varianza está relacionada con la varianza de la variable de partida mediante:

$$S_y^2 = e^2 S_x^2$$

En general, para cualquier momento centrado se tiene que: $m_r(y) = e^r m_r(x)$.

5. *Definición e interpretación de los coeficientes de asimetría y curtosis.*

Como indica su propia nombre, este coeficiente mide si los datos se reparten de igual forma un lado y otro de su posición central, expresada en este caso mediante la media.

$$g_1 = \frac{m_3}{S^3}$$

Se basa en que en distribuciones simétricas por cada observación a la derecha de la media hay otra a igual distancia a la izquierda, por tanto la expresión.

$$\sum_{i=1}^{n} \left(x_i - \bar{x} \right)^3 = 0$$

Si la gráfica es asimétrica a la izquierda se rompe el anterior equilibrio entre sumandos positivos y negativos, resultando la suma negativa. Lo mismo ocurre cuando la distribución de frecuencias es asimétrica a la derecha, resultando la suma positiva. El signo del denominador de g1 es siempre positivo, por tanto:

- Si la distribución es simétrica => g1 = 0
- Si la distribución es asimétrica a la izquierda => g1 < 0
- Si la distribución es asimétrica a la derecha => g1 > 0

Al expresarse m_3 y S^3 en las mismas unidades, su cociente es una medida adimensional que podrá utilizarse para comparar la asimetría de distintas distribuciones. Por ser g_1 adimensional no le afecta los cambios de escala, además los cambios de origen no afectan a los momentos centrados (en particular a S^2 y m_3), en consecuencia g_1 es independiente de cambios de origen y escala.

Coeficiente de curtosis

Las medidas de curtosis se utilizan en distribuciones unimodales simétricas o levemente asimétricas para cuantificar la mayor o menor frecuencia de observaciones en torno a la media. La mayor frecuencia de observaciones próximas a la media dará lugar a una representación gráfica más apuntada, la menor frecuencia de observaciones próximas a la media dará lugar a una representación más aplanada. El perfil de apuntamiento que se toma como referencia es el de la conocida campana de Gauss o curva normal. La medida de apuntamiento más utilizada es el coeficiente de curtosis de Fisher

$$g_2 = \frac{m_4}{S^4} - 3$$

Al igual que el coeficiente de asimetría de Fisher es adimensional e independiente de cambios de origen y escala. Los valores del coeficiente de curtosis de Fisher se interpretan de la siguiente manera:

- Si la distribución tiene un apuntamiento normal (mesocúrtica) => g2 = 0
- Si la distribución es más aplanada que la curva normal (platicúrtica); g2 < 0
- Si la distribución es más apuntada que la curva normal (leptocúrtica); g2 > 0

6. <u>¿Qué mide y cómo se construye la curva de Lorenz? Indique y comente las situaciones extremas que pueden presentarse.</u>

Mide gráficamente la mayor o menor concentración en el reparto de una cantidad. Para construirla es necesario tener todos los qi = Sumat(ui)100/Utotal y de pi = Sumat(Ni)100/Ntotal. Estando la qi en el eje X y la pi en el eje Y, se unen todos los puntos. A mayor área respecto de la bisectriz, mayor concentración. Si es igual que la bisectriz indica que está equidistribuida. Para darse una situación de máxima concentración en el pi = 100 pasa de tener un qi = 0 a qi = 100. Es decir, pinta todo el área desde la bisectriz al eje X.

<u>PREGUNTAS TEMA 2</u>

1. <u>Independencia de variables estadísticas.</u>

Se dice que X e Y son independientes estadísticamente si todas las distribuciones condicionadas coinciden, en otras palabras, las frecuencias relativas de las distribuciones condicionadas son iguales (además también coinciden con las frecuencias relativas marginales). Esto es lo mismo que decir que el comportamiento de X no depende del valor de Y y que el comportamiento de Y tampoco depende del valor de X.

Más cómodo, pero equivalente a comprobar que las frecuencias relativas de las distintas distribuciones condicionadas son iguales, es ver si se cumple

$$n_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n} \qquad \forall i, j$$

O equivalentemente $$n_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n} \Leftrightarrow \frac{n_{ij}}{n} = \frac{n_{i\bullet}}{n} \frac{n_{\bullet j}}{n} \Leftrightarrow f_{ij} = f_{i\bullet} f_{\bullet j} \qquad \forall i, j$$

2. *Coeficiente de correlación lineal. Interpretación.*

Su signo es el mismo que el de la covarianza, además este coeficiente toma valores entre -1 y 1; (-1 ≤ rxy ≤ 1) , midiendo objetivamente el grado de variación conjunta que tienen las variables X e Y. Cuando las variables están tipificadas el coeficiente de correlación lineal coincide con la covarianza. El cambio de origen y escala no afecta al valor de este coeficiente. Sabemos que

$$X^* = eX + c \qquad Y^* = hY + d \qquad \Rightarrow \qquad S_{x^*y^*} = ehS_{xy} \qquad S_{x^*} = eS_x \qquad S_{y^*} = hS_y$$

por tanto,

$$r_{x^*y^*} = \frac{S_{x^*y^*}}{S_{x^*} S_{y^*}} = \frac{ehS_{xy}}{eS_x hS_y} = \frac{S_{xy}}{S_x S_y} = r_{xy}$$

3. Si $r = 1$, existe una correlación positiva perfecta. El índice indica una dependencia total entre las dos variables denominada *relación directa*: cuando una de ellas aumenta, la otra también lo hace en proporción constante.
4. Si $0 < r < 1$, existe una correlación positiva.
5. Si $r = 0$, no existe relación lineal. Pero esto no necesariamente implica que las variables son independientes: pueden existir todavía relaciones no lineales entre las dos variables.
6. Si $-1 < r < 0$, existe una correlación negativa.
7. Si $r = -1$, existe una correlación negativa perfecta. El índice indica una dependencia total entre las dos variables llamada *relación inversa*: cuando una de ellas aumenta, la otra disminuye en proporción constante.

3. *Coeficiente de determinación. Interpretación.*

La media de los residuos o errores al cuadrado que se cometen en un ajuste se conoce como varianza residual y es una medida de la bondad del ajuste $$S_{ry}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - f(x_i)\right)^2$$

Esta medida tiene el inconveniente de no estar acotada entre unos valores fijos por lo que tenemos el problema de no saber con precisión si está tomando valores significativamente grandes o pequeños. Para superar dicho problema se define a partir de ella el coeficiente de determinación

como $$R^2 = 1 - \frac{S^{2l}_{ry}}{S^2_y} \qquad , siendo \quad 0 \le R^2 \le 1$$

En el caso de la recta, este coeficiente coincide con el coeficiente de correlación lineal al cuadrado.

Al no relacionarse (linealmente) la variable independiente con la variable dependiente y por tanto no aportar ninguna información, la variable independiente no aparece en la expresión de la recta de regresión que se reduce a estimar los valores de la variable Y mediante y sea quien sea el valor de X.

Si R^2=1 estamos ante un ajuste por mínimos cuadrados perfecto, la gráfica de la función pasa por todos los puntos de la nube (como en el ejemplo de vt=500). En el caso de la recta, como se dijo anteriormente, ambas rectas de regresión coinciden y pasan por todos los puntos. Para valores intermedios, 0 < R^2 < 1 , según esté R_2 más próximo a un extremo u otro nos indicará un peor o mejor ajuste. Además, en el caso de la recta, cuanto más se aproxime a 1 más se cierra el ángulo que forman las rectas de regresión y cuanto más se aproxime a cero mayor será dicho ángulo.

Otra interpretación de este coeficiente de determinación, R_2 , es la proporción de la varianza de Y (comportamiento de la variable Y) que puede atribuirse a su relación con X.

PREGUNTAS TEMA 5

1. *Definición de probabilidad. Asignación de probabilidades.*

La distinción entre fenómenos determinísticos y aleatorios conduce a plantearse el problema de la medida de la incertidumbre mostrada por estos últimos.
1.- La frecuencia relativa de cualquier modalidad x_i es un número comprendido entre 0 y 1

$$0 \le f_i \le 1.$$

2.- La frecuencia relativa de dos o más modalidades es la suma de las frecuencias relativas de cada una de las modalidades.

$$f(par) = f\left(x_2 \cup x_4 \cup x_6\right) = f_2 + f_4 + f_6$$

3.- La frecuencia relativa de todas las modalidades (suma de todas las frecuencias relativas) es 1.

$$\sum_{i=1}^{6} f_i = 1$$

Llamaremos suceso a cada uno de los posibles resultados de un experimento o fenómeno aleatorio. Utilizaremos letras mayúsculas para referirnos a ellos.

El suceso formado por un único resultado se denomina suceso elemental. Para estos sucesos, además de la notación general con letras mayúsculas, utilizaremos la notación ωi .
Se llama suceso seguro al suceso que siempre ocurre, está formado por todos los sucesos
elementales, se nota Ω.
Al conjunto vacío, que se nota ∅, se denomina suceso imposible.

2. Definición de probabilidad condicionada, Sucesos dependiente e independientes.

Se define la probabilidad del suceso B condicionado a A, ( ) B A P , como la probabilidad de que ocurra B supuesto que ha ocurrido A. Dicha probabilidad condicionada del suceso B al suceso A es igual a:

$$P\left(B/A\right) = \frac{P(B \cap A)}{P(A)} \qquad P(A) > 0$$

La definición de probabilidad condicionada, fijado A, verifica las condiciones de una medida de probabilidad:

1.- $P\left(B/A\right) \geq 0 \qquad \forall B \in \mathcal{A}$ **2.-** Sean $B_1, \ldots, B_n$ sucesos incompatibles, $P\left(\bigcup_{i=1}^{n} B_i / A\right) = \sum_{i=1}^{n} P\left(B_i/A\right)$

3.- $P\left(\Omega/A\right) = 1$

SUCESOS DEPENDIENTES E INDEPENDIENTES

En general $P\left(B/A\right) \neq P(B)$ y decimos en tal caso que *B* **depende** de *A*. Si $P\left(B/A\right) = P(B)$ diremos que *B* es **independiente** de *A*, es decir, que haya ocurrido *A* no modifica la probabilidad de *B*.

En el caso de **independencia:**

$$P\left(B/A\right) = \frac{P(B \cap A)}{P(A)} = P(B) \qquad \Rightarrow \qquad P(A \cap B) = P(A)P(B)$$

Y como consecuencia se sigue que también:

$$P\left(A/B\right) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$$

luego si *B* es independiente de *A*, entonces *A* es independiente de *B*.
Según todo lo anterior, se toma como definición de que los sucesos *A* y *B* son **independientes** la siguiente expresión:

$$P(A \cap B) = P(A)P(B)$$

La noción de independencia es fundamental en la teoría de la probabilidad. La mayoría de los resultados en probabilidad se obtienen bajo la hipótesis de independencia.

3. <u>*Fórmula de probabilidad total. Fórmula de Bayes.*</u>

Sean los sucesos 1, ..., A An una partición de Ω , es decir:

$$A_i \cap A_j = \varnothing \qquad \forall i \neq j \qquad ; \qquad \bigcup_{i=1}^{n} A_i = \Omega$$

en este contexto, la fórmula de la probabilidad total afirma que:

$$\mathrm{P}(B) = \sum_{i=1}^{n} \mathrm{P}\left(B/A_i\right) \mathrm{P}(A_i)$$

Demostración:

$$\mathrm{P}(B) = \mathrm{P}(B \cap \Omega) = \mathrm{P}\left(B \cap \left(\bigcup_{i=1}^{n} A_i\right)\right) = \mathrm{P}\left(\bigcup_{i=1}^{n} (B \cap A_i)\right) = \sum_{i=1}^{n} \mathrm{P}(B \cap A_i) = \sum_{i=1}^{n} \mathrm{P}\left(B/A_i\right) \mathrm{P}(A_i)$$

Las probabilidades condicionadas de un suceso cualquiera B sobre cada uno de los elementos de la partición,P(B/Ai) , son normalmente conocidas. No es así con las probabilidades(Ai/B) . Veamos cómo resolver este problema:

por la fórmula de la probabilidad compuesta $\quad \mathrm{P}(A_k \cap B) = \mathrm{P}(B)\mathrm{P}\left(A_k/B\right) = \mathrm{P}(A_k)\mathrm{P}\left(B/A_k\right)$

y según la fórmula de la probabilidad total $\quad \mathrm{P}(B) = \sum_{i=1}^{n} \mathrm{P}\left(B/A_i\right) \mathrm{P}(A_i)$

de ambas expresiones se sigue:

$$\mathrm{P}\left(A_k/B\right) = \frac{\mathrm{P}(A_k \cap B)}{\mathrm{P}(B)} = \frac{\mathrm{P}(A_k)\mathrm{P}\left(B/A_k\right)}{\sum_{i=1}^{n} \mathrm{P}\left(B/A_i\right) \mathrm{P}(A_i)}$$

Esta última expresión se conoce como **fórmula de Bayes**.