	1 VARIABLE CATEGORICA	Hipótesis	Supuestos	Estadístico de contraste	Distribución muestral	Zona crítica	SPSS
N P	**PRUEBA BINOMIAL.** Variables dicotómicas. Contraste sobre una proporción.	C.Bilateral: $H_0: \pi_1 = K_0; H_1: \pi_1 \neq K_0$ C.U.Dcho: $H_0: \pi_1 \leq K_0; H_1: \pi_1 > K_0$ C.U.Izdo: $H_0: \pi_1 \geq K_0; H_1: \pi_1 < K_0$ K_0 es la proporción que queremos contrastar.	1. Variables dicotómicas. 2. $\pi_1$es la proporción de éxitos en la población. 3. Se puede hacer corrección por continuidad si n es muy pequeño	n_1: numero de éxitos en los n ensayos. π_1: Proporción de éxitos en los n ensayos. Z	Binomial con parámetros n_1 y π_1. A medida que aumenta n se aproxima a una normal Z se aproxima a una normal N(0,1).	Bilateral: $Z \leq Z_{\alpha/2}$ o $Z \geq Z_{1-\alpha/2}$ Derecho: $Z \geq Z_{1-\alpha}$ Izquierdo: : $Z \leq Z_{\alpha}$ **Nivel crítico:** Bilateral: $p = 2[P(Z \geq \lvert Z_h \rvert)]$ Derecho: $p = P(Z \geq Z_h)$ Izquierdo: $p = P(Z \leq Z_h)$ Z_h es el valor concreto que toma el estadístico Z.	**Pruebas no paramétricas>Binomial.** Si la proporción es 0,5 spss considera contraste bilateral. Si la proporción de casos de la categoría de referencia es mayor que K_0 se considera contraste unilateral derecho.
N P	**Prueba X^2de PEARSON SOBRE BONDAD DE AJUSTE. Variables politómicas** Comprueban si la distribución teórica se ajusta a la distribución empírica.	$H_0: f(n_i) = M(n; \pi_1, \pi_2 \dots \pi_n)$ $H_1: f(n_i) \neq M(n; \pi_1, \pi_2 \dots \pi_n)$	Para variables categóricas y cuantitativas y En variables dicotómicas: Para ver el ajuste a la distribución binomial. En variables politómicas: para valorar ajuste a distribución multinomial. En variables cuantitativas. Para valorar ajuste a una distribución normal. 3. M.a de **N** observaciones en **I** categorías exclusivas. 4. Se asume **independencia** entre las observaciones. 5. No mas del 20% de las frecuencias esperadas es menor a 5	$X^2 = \sum_{i=1}^{I} \frac{(n_i - m_i)^2}{m_i}$	Se aproxima a una $\boldsymbol{X^2_{I-1}}$ grados de libertad.	$X^2 \geq X^2_{I-1;1-\alpha}$ **Nivel crítico** $p = P(X^2 \geq X^2_h)$	**Analizar> pruebas no paramétricas>chi cuadrado.**
	1 VARIABLE CUANTITATIVA.	**Hipótesis**	**Supuestos**	**Estadístico de contraste**	**Distribución muestral**	**Zona crítica**	**SPSS**
SOBRE EL CENTRO DE UNA DISTRIBUCIÓN	**CONTRASTE SOBRE UNA MEDIA. PRUEBA T** Comprueba si una media teórica se ajusta a la media empírica.	C.Bilateral: $H_0: \mu_1 = K_0; H_1: \mu_1 \neq K_0$ C.U.Dcho: $H_0: \mu_1 \leq K_0; H_1: \mu_1 > K_0$ C.U.Izdo: $H_0: \mu_1 \geq K_0; H_1: \mu_1 < K_0$ K_0 es el valor concreto de $\mu_1$que queremos contrastar.	1. Muestra aleatoria de tamaño n. 2. El supuesto de normalidad va perdiendo importancia conforme el tamaño muestral va aumentando. 3. Independencia: observaciones independientes. 4. Normalidad:	$T = \frac{Y - \mu_y}{S_y/\sqrt{n}}$	T_{n-1}grados de libertad	Bilateral: $T \leq t_{n-1;\alpha/2}$ y $T \geq t_{n-1;\alpha/2}$ Derecho: $T \geq t_{n-1;\,\alpha}$ Izquierdo: : $T \leq t_{n-1;\,\alpha}$ **Nivel Crítico (P valor)** Bilateral: $p = 2[P(T \geq T_h)]$ Derecho: $p = P(T \geq T_h)$ Izquierdo: $p = P(T \leq T_h)$ T_h es el valor concreto que toma el estadístico T.	**Analizar > Comparar medias > Prueba T para una muestra**
	PRUEBA DE WILCOXON. Permite constatar sobre el centro sin necesidad de asumir normalidad. Usa la mdn	Sobre la mediana. C.Bilateral: $H_0: Mdn_y = K_0; H_1: Mdn_y \neq K_0$ C.U.Dcho: $H_0: Mdn_y \leq K_0; H_1: Mdn_y > K_0$ C.U.Izdo: $H_0: Mdn_y \geq K_0; H_1: Mdn_y < K_0$ K_0 es el valor concreto de Mdn_yque queremos contrastar.	1. Distribución simétrica. (Inferencias efectuadas sobre la mdn en una distribución simétrica son trasladables a la media. 3. Muestra aleatoria de n observaciones. (Wilcoxon aprovecha info ordinal) Toma en consideración las distancias.	$S_+ = \sum R_{i(+)}$	Distribución muestral S_+	Bilateral: $S_+ < s_{\alpha/2}$ y $S_+ > s_{1-\alpha/2}$ Derecho $S_+ < s_{1-\alpha}$ Izquierdo: $S_+ > s_{\alpha}$ **Decisión** Se rechaza la H_0 si S_+cae en la zona critica.	*Hay que crear una variable con la categoría Mdn que se quiere comparar. Transformar> Calcular. **Analizar>Pruebas no Paramétricas> Muestras relacionadas> Wilcoxon**
	PRUEBA DE LOS SIGNOS o Binomial. Permite contrastar sobre el centros sin supuestos necesarios. Requiere una Y ordinal.	C.Bilateral: $H_0: Mdn_y = K_0; H_1: Mdn_y \neq K_0$ C.U.Dcho: $H_0: Mdn_y \leq K_0; H_1: Mdn_y > K_0$ C.U.Izdo: $H_0: Mdn_y \geq K_0; H_1: Mdn_y < K_0$ K_0 es el valor concreto de Mdn_yque queremos contrastar.	1. Muestra aleatoria 2. Inferencias efectuadas sobre la mdn. 3.Cada valor de la variable es clasificado como mayor, menor o igual que K_0para obtener $n_+, n_-, n_=$. (Aprovecha solo info nominal) (Solo tiene en cuenta el número de rangos positivos)	n_+: número de diferencias positivas. Z	n_+; según modelo de probabilidad binomial con parámetros n y $\pi_+ = 0{,}50$ Z: Distribución normal N(0,1)	Bilateral: se Rechaza H_0 si n_+ si la probabilidad de obtener ese valor es menor que α. Derecho: se Rechaza H_0 si n_+ si la probabilidad de obtener un valor como ese o mayor es menor que α. Izdo: se Rechaza H_0 si n_+ si la probabilidad de obtener un valor como ese o mas pequeño es menor que α	**Analizar>Pruebas no paramétricas> Binomial**
	CONTRASTE SOBRE UNA VARIANZA. Contraste sobre la dispersión.	C.Bilateral: $H_0: \sigma^2{}_y = K_0; H_1: \sigma^2{}_y \neq K_0$ C.U.Dcho: $H_0: \sigma^2{}_y \leq K_0; H_1: \sigma^2{}_y > K_0$ C.U.Izdo: $H_0: \sigma^2{}_y \geq K_0; H_1: \sigma^2{}_y < K_0$ K_0 es el valor concreto de $\sigma^2{}_y$ que interesa contrastar.	m.a. de tamaño n extraída de población normal.	$X^2 = (n-1)S_y/\sigma^2$	X^2_{n-1}	Biltl: $X^2 \leq X^2_{n-1;1-\alpha/2}$ y $X^2 \geq X^2_{n-1;1-\alpha/2}$ Derecho: $X^2 \geq X^2_{n-1;1-\alpha}$ Izquierdo: $X^2 \leq X^2_{n-1;1-\alpha}$	
	KOLMOGROV-SMIRNOV. Para estudiar la forma de la distribución. (Normalidad)	$H_0: F(Y_i) = F_0(Y_i); F(Y_i) \neq F_0(Y_i)$	m.a de n observaciones de una variable cuantitativa Y_i.	$D_{ks} = max\lvert F(Y_i) - F_0(Y_i) \rvert$.		$D_{ks} > d_{\alpha}$ **Nivel Crítico** $p = P(D_{ks} > k)$, donde k se refiere al valor concreto que toma el estadístico D_{ks}	**Analizar>descriptivos> Explorar>Gráficos con prueba de normalidad** **Analizar>Pruebas no paramétricas> K-S de una muestra.**

	2 VARIABLES CATEGÓRICAS	Hay que abordar dos tareas básicas: 1. Describir el comportamiento conjunto de ambas variables. 2. Averiguar si están relacionadas. (mediante la prueba de independencia de X^2 de PEARSON) Dos variables categóricas son independientes cuando el comportamiento de una de ellas no se ve alterado por la presencia de la otra o cuando las distribuciones condicionales de cualquiera de ellas son iguales en todas las categorías de la otra. Si existe relación hay que: 1. Cuantificar con alguna medida de asociación 2. Interpretarla a partir de los residuos típificados.					
		Hipótesis	**Supuestos**	**Estadístico de contraste**	**Distribución muestral**	**Zona crítica**	**SPSS**
	Prueba de X^2 de PEARSON sobre independencia. (tablas de contingencia)	H_0: X e Y son variables independientes. ($\pi_{ij} = \pi_{i+}\pi_{+j}$ para todo ij). H_1: X e Y no son variables independientes. ($\pi_{ij} \neq \pi_{i+}\pi_{+j}$ para todo ij).	m.a de n observaciones clasificada en las I x J combinaciones. La probabilidad de que una observación cualquiera pertenezca a cada una de las casillas se mantiene constante durante todo el proceso de clasificación. No más del 20% de las frecuencias esperadas son menores que 5.	$X^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\frac{(n_{ij}-m_{ij})^2}{m_{ij}}$ $0 \leq X^2 \leq \infty$	X^2 Se aproxima a $X^2_{(I-1)(J-1)}$ conforme n va aumentando.	$X^2 \geq X^2_{(I-1)(J-1);1-\alpha}$ **Nivel Crítico** $p = P(X^2 \geq X^2_h)$ donde X^2_hse refiere al valor muestral concreto que toma X^2	**Analizar>Estadísticos descriptivos>Tablas de contingencia. >Estadísticos> Chi cuadrados >casillas> Tipificados corregidos (residuos)**
	Mc Nemar. Contraste sobre la hipótesis de homogeneidad marginal en tablas 2x2. Hipótesis de homogeneidad marginal, es una hipótesis de simetría que permite valorar el cambio. Se contrasta la hipótesis de **simetría relativa** mediante la comparación de proporciones que son independientes entre si.	C.Bilateral:$H_0: \pi_{1+} = \pi_{+1}; H_1: \pi_{1+} \neq \pi_{+1}$ C.Dcho:$H_0: \pi_{1+} \leq \pi_{+1}; H_1: \pi_{1+} > \pi_{+1}$ C.Izdo:$H_0: \pi_{1+} \geq \pi_{+1}; H_1: \pi_{1+} < \pi_{+1}$	Proporciones relacionadas. Estudios longitudinales. Antes/después. m.a. de n sujetos en la que se ha metido una variable dicotómica en dos momentos distintos o dos variables dicotómicas con las mismas categorías; o bien, muestra aleatoria de n pares de sujetos en las que se ha mediado una variable dicotómica.	$\boldsymbol{X^2_{McNemar}} = \frac{(\lvert\boldsymbol{n_{12}} - \boldsymbol{n_{21}}\rvert - \mathbf{1})^2}{(\boldsymbol{n_{12}} - \boldsymbol{n_{21}})}$	$\boldsymbol{X^2_{McNemar}}$ se aproxima a una distribución ji-cuadrado con 1 grado de libertad X^2_1. La aproximación es buena incluso con muestras pequeñas.	Biltl: $X^2_{McNemar} \geq X^2_{1;1-\alpha}$ Derecho^2: $X^2_{McNemar} \geq X^2_{1;1-2\alpha}$ Izquierdo^2: $X^2_{McNemar} \geq X^2_{1;1-2\alpha}$ Se rechaza la H_0 si el estadístico del contraste cae en la zona crítica. **Nivel Crítico** Bilatl; $p = [P(X^2_1 \geq X^2_{McNemar})]$ Dcho; $p = 2[P(X^2_1 \geq X^2_{McNemar})]$ Izqd; $p = 2[P(X^2_1 \geq X^2_{McNemar})]$	**Pruebas no paramétricas>**si el número de cambios es ≤ 25 se usa: distribución binomial. **Estadísticos descriptivos>tablas de contingencias> (antes: filas)(después: columnas)> estadísticos>McNemar.**
			Simetría con Variables politómicas: la prueba de Bowker:	$\boldsymbol{X^2_{McNemar-bowker}} = \sum_{i<j}\sum\frac{(\boldsymbol{n_{ij}} - \boldsymbol{n_{ij}})^2}{(\boldsymbol{n_{ij}} - \boldsymbol{n_{ij}})}$	Distribución según la probabilidad ji –cuadrado con gl igual al número de casillas fuera de la diagonal principal de la tabla (i= filas, j= columnas)		**Pruebas no paramétricas>Dos muestras relacionadas>contrastar pares. Homogeneidad marginal. Tipo de prueba.**
	Medidas de asociación. Para cuantificar el grado de asociación. Estas medidas se basan en X^2 de PEARSON y son medidas nominales. Intentan cuantificar el grados de asociación aplicando algún tipo de corrección al estadístico de Pearson para hacerle tomar algún valor entre 0 y 1.	**Coeficiente de contingencia C de Pearson.**	Aprovechan información nominal. Intentan cuantificar el grado de asociación aplicando algún tipo de corrección al estadístico de Pearson para hacerle tomar un valor comprendido entre 0 y 1.	$\boldsymbol{C} = \sqrt{\boldsymbol{X^2}/(\boldsymbol{X^2}+\boldsymbol{n})}$ $\boldsymbol{C_{máx}} = \sqrt{(\boldsymbol{k}-\mathbf{1})/\boldsymbol{k}}$			**Analizar> estadísticos descriptivos>Tablas de contingencia**
		Coeficiente V de Cramer.		$\boldsymbol{V_{Cramér}} = \sqrt{\boldsymbol{X^2}/[\boldsymbol{n}(\boldsymbol{k}-\mathbf{1})]}$			**Analizar> estadísticos descriptivos>Tablas de contingencia**
		Coeficiente Phi	Para tablas 2x2	$\boldsymbol{\varphi} = \sqrt{\boldsymbol{X^2}/\boldsymbol{n}}$			**Analizar> estadísticos descriptivos>Tablas de contingencia**
		Residuos tipificados	Su distribución se aproxima a la N(0,1),. Sirven para interpretar con precisión el significado de la asociación detectada.				**Analizar> estadísticos descriptivos>Tablas de contingencia**
	Índices de riesgo. Entre una variable dicotómica FACTOR y otra variable dicotómica DESENLACE. Estudios longitudinales.	**Riesgo relativo.** Se compara lo que ocurre entre los sujetos expuestos y los que no están expuestos. Un riesgo relativo de 1 indica que la proporción de desenlaces es la misma en los dos grupos.					
	Odds Ratio. Cuantifica la relación entre dos variables dicotómicas.						
	Indices de acuerdo	**Kappa de cohen**. Para variables nominales.	**Gamma, tau-b y d.** Para variables ordinales.				

	1 V. Categórica y 1 variable cuantitativa	Hipótesis	Supuestos	Estadístico de contraste	Distribución muestral	Zona crítica	SPSS
P	**PRUEBA T PARA DOS MUESTRAS INDEPENDIENTES.** Una variable categórica que define **2 grupos** y una cuantitativa.	C.Bilateral: $H_0: \mu_{y_1} - \mu_{y_2} = K_0$; $H_1: \mu_{y_1} - \mu_{y_2} \neq K_0$ C.Dcho: $H_0: \mu_{y_1} \leq \mu_{y_2} = K_0$; $H_1: \mu_{y_1} - \mu_{y_2} > K_0$ C.Izdo: $H_0: \mu_{y_1} - \mu_{y_2} \geq K_0$; $H_1: \mu_{y_1} - \mu_{y_2} < K_0$	Dos muestras de tamaños n_1 y n_2 seleccionadas aleatoriamente e independientemente. Homocedasticidad. Si no se da: $n_1 = n_2$ y supuesto de normalidad Supuesto de normalidad. Si no: N>20-25 N<20 distribuciones simétricas. Variable cuantitativa de intervalo o razón	T	T se ditribuye según t_{gl} con gl=$n_1 + n_2 - 2$	Biltl: $T \leq T_{gl;\alpha/2}$ y $T \geq T_{gl;1-\alpha/2}$ Derecho: $T \geq T_{gl;1-\alpha}$ Izquierdo: $T \leq T_{gl;1-\alpha}$ **Nivel crítico** Bilatl; $p = 2[P(T \geq \|T_h\|)]$ Dcho: $p = P(T \geq \|T_h\|)$ Izqd; $p = P(T \leq \|T_h\|)$ Siendo T_h el valor concreto que toma el estadístico T	**Analizar>Prueba T> Prueba T muestras independientes**
N P	**U MANN WHITNEY. MUESTRAS INDEPENDIENTES.** Una variable categórica que define **2 grupos** y una cuantitativa.	C.Bilateral: $H_0: E(Y_1) = E(Y_2)$; $H_1: E(Y_1) \neq E(Y_2)$. C.Dcho: $H_0: E(Y_1) \leq E(Y_2)$; $H_1: E(Y_1) > E(Y_2)$. C.Izdo: $H_0: E(Y_1) \geq E(Y_2)$; $H_1: E(Y_1) < E(Y_2)$.	Una variable al menos es **ordinal** medida en dos muestras de tamaños n_1 y n_2 seleccionadas aleatoria e independientemente de dos poblaciones con la misma forma: $F(Y_1) = F(Y_2)$	$U = S_1 = \sum_{i=1}^{n_1} R_{i1}$ $Z = \frac{U - n_1(N+1)/2}{\sqrt{n_1 n_2 (N+1)/12}}$	Distribución según U. La tabla incluye los puntos críticos de u_α. Para los puntos críticos de las colas de la derecha: $u_{1-\alpha} = n_1(N+1) - u_\alpha$ La distribución Z se aproxima a N(0,1) a medida que los tamaños muestrales van aumentando.	Biltl: $U < u_{\alpha/2}$ y $U > u_{\alpha/2}$ $Z \leq Z_{\alpha/2}$ y: $U \geq u_\alpha$ Derecho: $U > u_{1-\alpha}$ $Z \geq u_{1-\alpha}$ Izquierdo: $U < u_\alpha$ $Z \leq u_\alpha$ Rechazar si el estadístico de contraste cae en la zona crítica. **Nivel Crítico** Bilatl; $p = 2[P(Z \geq \|Z_h\|)]$ Dcho: $p = P(Z \geq \|Z_h\|)$ Izqd; $p = P(Z \leq \|Z_h\|)$ Siendo Z_h el valor concreto que toma el estadístico Z	**Pruebas no paramétricas>dos muestras independientes>Contrastar variables>Variable de agrupación.** **Definir grupos** introducir códigos 1 y 2 (de la variable grupo) **Opciones> descriptivos> continuar.**
N P	**K NIVELES. KRUSKAL WALLIS MUESTRAS INDEPENDIENTES.** Una variable categórica que define K grupos y una cuantitativa.	H_0: las J distribuciones poblacionales son iguales. H_1: Las J distribuciones no tienen la misma media.	Variable al menos ordinal medida en J muestras aleatoria e independientes extraídas de sus respectivas poblaciones, se asume que tienen la misma **forma**. No necesita establecer supuesto sobre las poblaciones originales. Permite trabajar con datos ordinales. Potencia limitada vs ANOVA I Factor. Lógica similar a U de Mann-Whitney.	$H = \frac{12}{N(n+1)} \sum_{j=1}^{J} \frac{R_j^2}{n_j} - 3(N+1)$	Se distribuye según el modelo de probabilidad *ji*-cuadrado, con J-1 gl.	$H \geq h_{1-\alpha}$ o $H \geq X^2_{J-1:1-\alpha}$ **Nivel Crítico** P=P(H> H_h) Donde H_h se refiere al valor concreto de H.	**Analizar> pruebas no paramétricas>para K muestras independientes. Variable de agrupación> definir grupos.**
	Medidas del tamaño del efecto.	La significación estadística no expresa lo fuerte o intensa que es la relación que delata una diferencia significativa.	**Para Dos grupos:** δ de Cohen Correlación de Pearson R_{XY} Medida de lenguaje común. **Para mas de dos grupos:** Eta-cuadrado. **Correcciones de η^2:** $\eta^2_{corregida}$ épsilo-cuadrado. Omega-cuadrado **δ de Cohen** La relevancia del tamaño del efecto se justifica desde el contexto. Aunque Cohen da unas referencias: *Grande d* ≈ 0,8 /0,40 *Mediano d* ≈ 0,5/ 0,25 *Pequeño d* ≈ 0,2 /0,10 Es más utilizada y es una medida tipificada. Basada en la diferencia de medias.	**R_{XY}** Elevando R_{XY} se obtiene la proporción de varianza que comparten ambas variables. (Coeficiente de determinación) Prestar atención al contexto. Referencias: *Grande r* ≈ 0,5 *Mediano r* ≈ 0,3 *Pequeño r* ≈ 0,1 Tiene un mínimo y un máximo.	**Medida de lenguaje común** No requiere asumir poblaciones normales ni varianzas iguales. Se puede usar con variables ordinales. Estima el numero de veces que cada puntuación Y_1 es mayor que cada puntuación Y_2 y dividiendo este recuento entre el numero total de comparaciones.	**Eta-cuadrado η^2** Expresa el grado de *asociación* entre la variable categórica y la variable cuantitativa. Se puede interpretar como proporción de varianza común, es decir, como el grado en que aumenta nuestro conocimiento de la variable cuantitativa por saber a que grupo pertenecen. Cohen siguiere: *Grande eta* ≈ 0,14 *Mediano eta* ≈ 0,06 *Pequeño eta* ≈ 0,01	**Épsilon cuadrado ε^2** Corrige ligeramente el sesgo de η^2 **Omega-cuadrado ω^2** Es de las más aceptadas. Su valor concreto depende de que sea de efectos fijo o aleatorios.
	Cálculo de la potencia del tamaño muestral.	Se obtiene la potencia de un contraste basado en la prueba T a partir de una sencilla transformación	δ = tamaño del efecto tipificado. $\varphi = \delta\sqrt{n}$	**Calcular a priori la potencia del contraste** Para conocer el tamaño muestral que necesitamos despejamos n. $n = \varphi^2/\delta^2$	**Lambda λ** Potencia de contraste de F (ANOVA) **Parámetro de no centralidad.** Indica como de grande es la variabilidad de las medias. Su transformación: $\varphi = \sqrt{\lambda/J}$		

	2 VV cuantitativas.	Para el análisis de dos variables cuantitativas se puede hacer con los datos de un mismo sujeto en distinto momento (diseños intrasujeto) o utilizar pares de sujetos. En cualquier caso las puntuaciones están relacionadas. La comparación de dos variables cuantitativas tiene sentido ante variables con la misma métrica, pero siempre se pueden relacionar independientemente de la métrica en la que se encuentren.					
		Hipótesis	**Supuestos**	**Estadístico de contraste**	**Distribución muestral**	**Zona crítica**	**SPSS**
	Prueba T para dos muestras relacionadas.	C.Bilateral:$H_0: \mu_D = K_0$; $H_1: \mu_D \neq K_0$ C.Dcho:$H_0: \mu_D \leq K_0$; $H_1: \mu_D > K_0$ C.Izdo:$H_0: \mu_D \geq K_0$; $H_1: \mu_D < K_0$	muestra aleatoria de n diferencias procedentes de una población normal. (el supuesto de normalidad pierde importancia conforme el tamaño muestral aumenta)	$T = \frac{\bar{Y} - \mu_D}{S_D/\sqrt{n}}$	Se distribuye como según una t con n-1 gl (T_{n-1})	Bilateral: $T \leq t_{n-1;\alpha/2}$ y $T \geq t_{n-1;\alpha/2}$ Derecho: $T \geq t_{n-1;\,\alpha}$ Izquierdo: : $T \leq t_{n-1;\,\alpha}$ **Nivel crítico** Bilateral: $p = 2[P(T \geq T_h)]$ Derecho: $p = P(T \geq T_h)$ Izquierdo: $p = P(T \leq T_h)$ T_h es el valor concreto que toma el estadístico T.	**Analizar>comparar medias>prueba T para muestras relacionadas.**
NP	**WILCOXON.** Permite comparar los centros de dos variables cuantitativas sin necesidad de normalidad.	C.Bilateral: $H_0: Mdn_{y1} = Mdn_{y2}$; $H_1: Mdn_{y1} = Mdn_{y2}$ C.U.Dcho: $H_0: Mdn_{y1} \leq Mdn_{y2}$; $H_1: Mdn_{y1} > Mdn_{y2}$ C.U.Izdo: $H_0: Mdn_{y1} \geq Mdn_{y2}$; $H_1: Mdn_{y1} < Mdn_{y2}$	Asume asimetría. Exige medidas de intervalo o de razón y aprovecha información ordinal. Se calcula con las diferencias, desechando las diferencias nulas y a las diferencias no nulas se le asignan rangos de 1 a n. Si las mdn poblacional son iguales habrá tantos valores $Y_1 > Y_2$	$S_+ = \sum R_{i(+)}$	Se distribuye según S_+	Bilateral: $S_+ < s_{\alpha/2}$ y $S_+ > s_{1-\alpha/2}$ Derecho $S_+ < s_\alpha$ Izquierdo: $S_+ > s_{1-\alpha}$ **Decisión** Se rechaza la H_0 si S_+cae en la zona critica.	**Analizar>Pruebas no paramétricas>muestras relacionadas>Wilcoxon.**
NP	**Prueba de Signos**	C.Bilateral: $H_0: Mdn_{y1} = Mdn_{y2}$; $H_1: Mdn_{y1} \neq Mdn_{y2}$; C.U.Dcho: $H_0: H_0: Mdn_{y1} \leq Mdn_{y2}$; $H_1: Mdn_{y1} > Mdn_{y2}$; C.U.Izdo: $H_0: Mdn_{y1} \geq Mdn_{y2}$; $H_1: Mdn_{y1} < Mdn_{y2}$;	1. Muestra aleatoria 2. Inferencias efectuadas sobre la mdn. 3.Cada valor de la variable es clasificado como mayor, menor o igual para obtener $n_+, n_-, n_=$. 4. No asume asimetría. (Aprovecha solo info nominal) (Solo tiene en cuenta el número de rangos positivos)	n_+: número de diferencias positivas. Z	n_+; según modelo de probabilidad binomial con parámetros n y $\pi_+ = 0{,}50$ Z: Distribución normal N(0,1)	Bilateral: se Rechaza H_0 si n_+ si la probabilidad de obtener ese valor es menor que α. Derecho: se Rechaza H_0 si n_+ si la probabilidad de obtener un valor como ese o mayor es menor que α. Izdo: se Rechaza H_0 si n_+ si la probabilidad de obtener un valor como ese o mas pequeño es menor que α	**Analizar>Pruebas no paramétricas> Binomial**
	Coeficiente de correlación de Pearson.	Cuantifica la intensidad de la relación entre dos variables y el sentido de la misma (+,-) El grado en que la covarianza alcanza su máximo. Sensible a la presencia de casos anómalos. Se asume distribución normal. Variables analizadas sean de intervalo o de razón.		$R_{XY} = \frac{S_{XY}}{S_Y\, S_X}$			**Analizar>correlaciones>Bivariadas**
	Coeficiente de correlación de SPEARMAN	Bilateral: H_0: X e Y no están relacionadas. H_1: la relación entre X e Y es monótona. Derecho: H_0: X e Y no están relacionadas. H_1: la relación entre X e Y es monótona creciente Izquierdo: H_0: X e Y no están relacionadas. H_1: la relación entre X e Y es monótona decreciente	Estudia la relación entre variables ordinales. Transformar las puntuaciones originales en rangos. Describe relaciones monótonas. Requiere medidas independientes.	$R_S = 1 - \frac{6\sum D_i^2}{n^3 - \mathrm{n}}$ $T = \frac{R_S\sqrt{n-2}}{\sqrt{1-R_S^2}}$	$r_{1-\alpha}$ para n≤ 30. T se aproxima al modelo de probabilidad de t de student con n-2 gl. T_{n-2}	Bilateral: $R_S < r_{\alpha/2}$ y $R_S > r_{1-\alpha/2}$ $T < t_{n-2;\,\alpha/2}$ y $T > t_{n-2;\,1-\alpha/2}$ Derecho $R_S > r_{1-\alpha}$ $T \geq t_{n-2;\,1-\alpha/2}$ Izquierdo: $R_S < r_\alpha$ $T \leq t_{n-2;\,\alpha/2}$	**Analizar>correlaciones>Bivariadas**
	Correlación Parcial	Expresa el grado de relación lineal entre dos variables. Es una técnica de control estadístico que permite cuantificar la relación neta entre dos variables al eliminar de ambas el efecto de terceras variables. Correlación de primer orden: se está controlando el efecto de una sola variable. Correlación de 2º Orden: Se esta controlando el efecto de dos variables.					**Analizar>Variables Controlando para** (variables cuyo efecto se desea controlar) **Opciones>Estadísticos> correlaciones de orden cero**

ANOVA I FACTOR						
	Hipótesis	**Supuestos**	**Estadístico de contraste**	**Distribución muestral**	**Zona crítica**	**SPSS**
ANOVA I FACTOR Completamente aleatorizado. Permite comparar mas de 2 variables. 1 Variable categórica (ordinal o nominal) con k grupos (**VI**) y 1 variable cuantitativa (**VD**)	Se basa en la variabilidad de las muestras. *Variabilidad intergrupos* refleja las diferencias entra las puntuaciones de cada muestra. *Variabilidad intergrupos* refleja la variabilidad entre las muestras. Al estimador basado en la *variabilidad intragrupos* se le llama **media cuadrática intragrupos** $\boldsymbol{MC_E}$ (media cuadrática error o residual. Independiente de las medias poblacionales) Al estimador basado en la *variabilidad intergrupos* se le llama **media cuadrática intergrupos** $\boldsymbol{MC_I}$ (dependiente de las medias poblacionales). El tamaño relativo de $\boldsymbol{MC_I}$ respecto del de $\boldsymbol{MC_E}$ está informando del grado de parecido existente entre las medias poblacionales **= F.**					**Dos procedimientos:** **1.Analizar>comparar medias> Anova de un factor.** (chequea igualdad de varianzas con la prueba de Levene y contrasta con los estadísticos de **Welch y Brown Fortsythe** cuando no hay varianzas iguales. **2. Analizar> modelo lineal general>univariante.** (incluye; η^2, potencia observada λ, comparaciones planeadas)
	H_0: las J medias poblacionales son iguales. H_1: Las J medias poblacionales no tienen la misma media.	**Independencia** entre las puntaciones. (Prueba de las rachas) **Normalidad** de las poblaciones. Si no se cumple usar Kruskal-wallis **Homocedasticidad** si no se cumple usar las correcciones de Brown y Forsythe o Welch. J muestras *aleatoriamente seleccionadas* de J poblaciones *normales* con la *misma varianza*.	$F = \frac{MC_{INTER}}{MC_E}$	$F_{(J-1)(N-J)}$	$F \geq F_{(J-1)(N-J);1-\alpha}$ **Nivel Crítico** $p = P(F \geq F_h)$ Siendo F_h el valor muestral concreto que toma el estadístico F	
COMPARACIONES MULTIPLES. Para concretas que medias difieren de que otras.	**Comparaciones planeadas o a priori.** Comparaciones lineales es una combinación lineal de medias con pesos o coeficientes, no todos iguales a 0, que suman 0. **Total posibles comparaciones:** $J(J-1)/2$	*Comparaciones entre pares* : $\boldsymbol{\psi_1 = (\mu_1 - \mu_2)}$ *Comparaciones múltiples:* $\boldsymbol{\psi_1 = (\mu_1 + \mu_2) - (\mu_3 + \mu_4)}$ Para realizar las comparaciones adecuadas hay que asignar **coeficientes** a cada media. **Tasa de error** Hacer múltiples comparaciones esta asociado a cometer mayor probabilidad de errores tipo I $P = 1 - (1-\alpha)^k$	**Prueba de Dunn-Bonferroni** Se trata del estadístico T aplicando una estrategia para controlar la tasa de error. Se va haciendo más conservadora y perdiendo potencia conforme aumenta el número de comparaciones.	**Comparaciones de tendencia** Sobre el tipo de relacione que tienen las variables. Es necesario que estén ordenadas. Lineal: con 2 o más VV. Cuadrática: con 3 o más VV. Cúbica: con 4 o más VV.	**Prueba de Dunnett** Controla la tasa de error cuando se realizan las J-1 comparaciones entre los grupos. Consiste en obtener un valor llamado DMS= diferencia mínima significativa, que es el valor más pequeño a partir del cual una diferencia es significativa.	**ANOVA I Factor> Contrastes>Polinómico>Orden = cuadrático** **Coeficientes>** Ir añadiendo los contrastes con sus coeficientes.
	Comparaciones Post-Hoc	**Prueba de Tukey** Diseñada para controlar la tasa de error cuando se llevan a cabo las $J(J-1)/2$ comparaciones. Basada en la T de Student. Utiliza la DMS: Diferencia mínima significativa.	**Prueba de Dunnet** Permite hacer comparaciones con un grupo en concreto. J-1	**Prueba de Scheffé** Controla la tasa de error para el total de posibles comparaciones. Es un procedimiento conservador.	**Cuando no hay varianzas iguales:** Games-Howell (sustituyendo a Tukey) Brown-Forsythe (sustituyendo a Scheffé)	**Comparar medias>ANOVA un factor>Post Hoc>Tukey.**
ANOVA I FACTOR. Medidas repetidas. Se utilizan los mismos sujetos en más de una condición. (Bloques aleatorios). Un solo grupo de sujetos por cuyos niveles pasan todos los sujetos.	La variabilidad total recoge la variabilidad entre cada observación y la media total. Se puede descomponer en 3 fuentes de variabilidad: 1. *Variabilidad intergrupos*: variabilidad de las puntuaciones de los mismos sujetos (intrasujeto). **Media cuadrática intergrupos** $\boldsymbol{MC_I}$ 2. *Variabilidad intersujetos:* la que se da entre los diferentes sujetos. **Media cuadrática intersujetos** $\boldsymbol{MC_S}$ 3. *Variabilidad error:* La que se da entre cada observación y sus medias marginales. (Distancia a la media) **Media cuadrática error** $\boldsymbol{MC_E}$					**Analizar>modelo lineal general>medidas repetidas.** **Introducir el nombre del FACTOR>añadir numero de niveles>Definir.** **Variables intrasujetos** **Factores e interacciones de los factores> mostrar las medias.** **Comparar los efectos principales>bonferroni>Ajuste del intervalo de confianza>comparar efectos principales.**
	H_0: las J medias poblacionales son iguales. H_1: Las J medias poblacionales no tienen la misma media.	**Independencia:** Entre los sujetos **Normalidad:** **Homocedasticidad:** **Esfericidad:** Las diferencias de las varianzas son iguales. Es la condición necesaria para que F se distribuya con los gl propuestos. (la matriz de varianzas-covarianzas debe ser esférica= **Simetría compuesta.Prueba de Mauchly.**	$F = \frac{MC_{INTER}}{MC_{ERROR}}$	$F_{J-1,(n-1)(J-1)}$	$F \geq F_{J-1,(n-1)(J-1):1-\alpha}$ **Nivel Critico** $p = P(F \geq F_h)$ Siendo F_h el valor muestral concreto que toma el estadístico F	
	Si no hay esfericidad: El estadístico F se vuelve liberal. Hay que elegir otras opciones, aumenta la probabilidad de cometer errores tipo I.	**Modificar los gl de la distribución F** $F_{\varepsilon(J-1),\varepsilon\,(n-1)(J-1)}$ (Geiser y Greenhouse y Huynh y Feldt)	**Estadísticos multivariados:** Traza de Pillai Lambda de Wilks Traza de Hotelling Raíz Mayor de Roy	**Orden de preferencia** 1. Si hay esfericidad ⇒F 2. Si no hay esfericidad ⇒Estadísticos multivariados. Si se rechaza la H_0: ⇒ FIN Si no se rechaza la H_0: ⇒ gl corregidos.		
NP **FRIEDMAN**		Para datos ordinales, solo aprovecha información ordinal. No requiere normalidad ni esfericidad. Es una extensión de Wilcoxon.	X_r^2			**Analizar>pruebas no paramétricas> K muestras relacionadas>contrastar variables.**
NP **PRUEBA DE COCHRAN**		Para variables dicotómicas.				**Analizar>pruebas no paramétricas> K muestras relacionadas>contrastar variables> tipo de prueba> Cochran**

<table>
<tr><td colspan="7">ANOVA II FACTORES</td></tr>
<tr><td></td><td>**Hipótesis**</td><td>**Supuestos**</td><td>**Estadístico de contraste**</td><td>**Distribución muestral**</td><td>**Zona crítica**</td><td>**SPSS**</td></tr>
<tr><td rowspan="4">**ANOVA II FACTORES.**
Dos variables independientes categóricas y una variable dependiente cuantitativa.</td><td colspan="5">Los efectos que interesa estudiar son:
Los efectos principales: los efectos de cada factor individualmente considerado
Los efectos simples: el efecto del factor teniendo en cuenta un único nivel del otro factor.
El efecto de la interacción: es el efecto conjunto de los dos factores. Hay interacción cuando el efecto de uno de ellos sobre la VD no es el mismo en todos los niveles del otro factor.
El resultado de la combinación de dos factores difiere de la suma de los efectos principales de esos factores. Las diferencias entre cada uno de los efectos simples es interacción.
(*efectos principales, efecto de la interacción)*
Media cuadrática intragrupos MC_E (media cuadrática error o residual. Independiente de las medias poblacionales)
Media cuadrática intergrupos MC_I (dependiente de las medias poblacionales). Se halla una media cuadrática por factor.
El tamaño relativo de MC_I respecto del de MC_E está informando del grado de parecido existente entre las medias poblacionales **= F.**</td><td rowspan="4">**Modelo lineal general> univariante Dependiente y factores fijos**

Opciones> tamaño del efecto y potencia observada.

Efectos principales:
Post hoc> Tukey y Games-Howell.

Opciones>homogeneidad. (levene)

***Efectos simples:* Opciones> mostrar las medias para>comparar los efectos principales> Bonferroni>Ajuste del intervalo de confianza PEGAR** (sintaxis; en la línea EMMEANS=TABLES(factorA*Vdependiente) **añadir** COMPARE(FactorB)ADJ (BONFERRONI)

Efecto de la interacción:
1. Gráfico.
2. Comparar entre si los efectos simples.
Comparar medias>ANOVA de un factor>VD a Dependientes> *Grupo a Factor**
Contrastes>introducir coeficientes.
Grupo* variable que creamos con los niveles de las VVII 3x2.</td></tr>
<tr><td rowspan="3">a. $H_{0(A)}$: las medias poblacionales correspondientes con las J niveles del factor son iguales.
$H_{1(A)}$: las medias poblacionales correspondientes con las J niveles del factor no son iguales.

b. $H_{0(B)}$: las medias poblacionales correspondientes con las K niveles del factor son iguales.
$H_{1(B)}$: las medias poblacionales correspondientes con las K niveles del factor no son iguales.

c. $H_{0(AB)}$: No hay efecto de la interacción.
$H_{1(AB)}$: Hay efecto de la interacción.</td><td>**Independencia.**
Normalidad
Homocedasticidad.</td><td>$F_A = \frac{MC_{INTER(A)}}{MC_{ERROR}}$

$F_B = \frac{MC_{INTER(B)}}{MC_{ERROR}}$

$F_{AB} = \frac{MC_{INTER(AB)}}{MC_{ERROR}}$</td><td>$F_A$ se distribuye $F_{(J-1),(N-JK)}$

F_B se distribuye $F_{(K-1),(N-JK)}$

F_{AB} se distribuye $F_{(J-1)(K-1),(N-JK)}$</td><td>a. $F_A \geq F_{(J-1),(n-JK);1-\alpha}$

b. $F_B \geq F_{(K-1),(n-JK);1-\alpha}$

c. $F_{AB} \geq F_{(J-1)(K-1),(N-JK;1-\alpha}$</td></tr>
<tr><td>**COMPARACIONES MÚLTIPLES**</td><td>**Efectos principales***</td><td>**Efectos simples**</td><td>**Efecto de la interacción**</td></tr>
<tr><td></td><td>**A priori**
Dunn Bonferroni.
Prueba de Dunnett
Post Hoc:
Tukey
Scheffé</td><td>Para entender el significado de la interacción tenemos que atender a las diferencias entre los efectos simples.</td><td>Se hacen las comparaciones ordenando los coeficientes, para comparar si las diferencias entre los efectos simples son significativas
$(\mu_{11} - \mu_{21}) - (\mu_{12} - \mu_{22})$

$(1)\mu_{11} + (-1)\mu_{21} + (-1)\mu_{12}(+1\mu_{22})$</td></tr>
<tr><td>**ANOVA II FACTORES. Medidas repetidas en ambos factores.**
Dos variables independientes categóricas y una variable dependiente cuantitativa.
Bloques de sujetos.
Mismos sujetos.</td><td>a. $H_{0(A)}$: las medias poblacionales correspondientes con las J niveles del factor son iguales.
$H_{1(A)}$: las medias poblacionales correspondientes con las J niveles del factor no son iguales.

b. $H_{0(B)}$: las medias poblacionales correspondientes con las K niveles del factor son iguales.
$H_{1(B)}$: las medias poblacionales correspondientes con las K niveles del factor no son iguales.

c. $H_{0(AB)}$: No hay efecto de la interacción.
$H_{1(AB)}$: Hay efecto de la interacción.</td><td>**Independencia:** Entre los sujetos
Normalidad:
Homocedasticidad:
Esfericidad:
Esfericidad local: de cada factor independientemente.</td><td>$F_A = \frac{MC_{INTER(A)}}{MC_{ERROR(AxS)}}$

$F_B = \frac{MC_{INTER(B)}}{MC_{ERROR(BxS)}}$

$F_{AB} = \frac{MC_{INTER(AB)}}{MC_{ERROR(ABxS)}}$</td><td>$F_A$ se distribuye $F_{J-1,(J-1)\,(n-1)}$

F_B se distribuye $F_{K-1,(K-1)(n-1)}$

F_{AB} se distribuye $F_{(J-1)(K-1),(J-1)(K-1)(n-1)}$</td><td>a. $F_A \geq F_{J-1,(J-1)\,(n-1);1-\alpha}$

b. $F_B \geq F_{K-1,(K-1)(n-1);1-\alpha}$

$c.\ F_{AB} \geq F_{(J-1)(K-1),(J-1)(K-1)(n-1);1-\alpha}$</td><td>**Modelos lineal general> medidas repetidas> Asignar los dos factores.**

Variables intrasujetos

Gráficos

Opciones>mostrar las medias para>comparar efectos principales> Bonferroni >ajuste del intervalo de confianza.

Opciones>Estadísticos descriptivos>Estimaciones del tamaño del efecto>potencia observada.

***Efectos simples:* Contrastes>simples >Primera> cambiar.**</td></tr>
<tr><td>**ANOVA II FACTORES. Medidas repetidas en un factor.**</td><td>Las mismas que en ANOVA II-CA</td><td>***Efecto intersujetos:***
Independencia: Entre los sujetos
Normalidad:
Homocedasticidad:
Efecto intrasujetos:
Esfericidad
Esfericidad multi-muestra (Prueba de Mauchly)
Igualdad de matrices (Prueba de Box)
Si no hay esfericidad usar las alternativas al estadístico F.</td><td>$F_A = \frac{MC_{INTER(A)}}{MC_{ERROR(S)}}$

$F_B = \frac{MC_{INTER(B)}}{MC_{ERROR(BxS)}}$

$F_{AB} = \frac{MC_{INTER(AB)}}{MC_{ERROR(BxS)}}$</td><td>$F_A$ se distribuye $F_{(J-1)\,,J(n-1)}$

F_B se distribuye $F_{K-1,J(K-1)(n-1)}$

F_{AB} se distribuye $F_{(J-1)(K-1),J(K-1)(n-1)}$</td><td>a. $F_A \geq F_{(J-1),J\,(n-1);1-\alpha}$

b. $F_B \geq F_{K-1,J(K-1)(n-1);1-\alpha}$

$c.\ F_{AB} \geq F_{(J-1)(K-1),J(K-1)(n-1);1-\alpha}$</td><td>**Modelo lineal general> medidas repetidas>Nombre del Factor intrasujetos >Añadir> definir las VV. Factores intersujetos.**

Gráficos

Post Hoc:* Opciones> mostrar las medias para>comparar los efectos principales>Bonferroni>Ajuste del intervalo de confianza>*Sintaxis..

***Efectos simples*: Contrastes> simple**</td></tr>
</table>

REGRESIÓN LINEAL SIMPLE. Dos variables. Explora y cuantifica la relación entre una variable cuantitativa VD y otra cuantitativa o categórica VI.	**Para describir dos variables:**	**Recta de regresión**:	**Coeficientes de regresión tipificados:**	**Bondad de ajuste:**	**SPSS**
	1. Forma de la nube de puntos. Si hay alguna tendencia. 2. Centro: resumir la nube de puntos en una línea. 3. Valorar el grado de concentración o dispersión.	$Y = B_0 + B_1X$ Hay que averiguar cual es la recta que mejor define a la nube de puntos, se hace mediante **mínimos cuadrados** (suma de los cuadrados de las distancias verticales entre cada punto y la recta) Elegimos la recta que hace mínimas las distancias verticales.	El signo de B_1 indica si la relación es positiva o negativa, indica el cambio esperado en y por el aumento de X. $(Pronostico)Z_{VD} = Z_{VI}X$	Es importante evaluar el grado de ajuste de esa recta a la nube de puntos. $\boldsymbol{R^2_{XY}}$ cuantifica la relación. Es el % de varianza común explicada. Se puede interpretar como **proporción de reducción de los errores de predicción.**	**Analizar>regresión lineal**
		Coeficientes de regresión tipificados:			
		El signo de B_1 indica si la relación es positiva o negativa, indica el cambio esperado en y por el aumento de X.			
REGRESIÓN LINEAL MULTIPLE. Con más de una variable independiente.		**Ecuación de regresión**	$(Pronostico)Z_{VD} = Z_{VI_1}X_1 + Z_{VI_2}X_2 \dots Z_{VI_p}X_p$	Para saber si la ecuación es adecuada utilizamos el **coeficiente de correlación múltiple.**	**Regresión>lineal>** Trasladar **VD** y **VVII**. **Estadísticos> intervalos de confianza>Coeficientes de regresión.** *Pronósticos*: **Guardar> No tipificados> valores pronosticados> media** e **individuos (intervalos de pronósticos).** ***Importancia relativa de las variables:*** **Estadísticos> Correlaciones parciales y semiparciales** (las semiparciales indican el grado de relación VI-VD tras eliminar el efecto atribuible al resto de VI). ***Supuestos***: ***Colinealidad:*** **Gráficos>generar gráficos parciales** **Estadísticos> diagnósticos de colinealidad.** ***Distribución:*** **> histograma> grafico de los residuos tipificados> gráficos de probabilidad normal.** ***Independencia:*** **Durbin Watson** (1,5-2,5) ***Homocedasticidad:*** **Graficos>*ZPRED** al eje de las **X** y la variable ***ZRESID** al eje de las **Y.** ***Regresión jerárquica:*** **>Método> hacia delante.** **>estadísticos>Cambio en R cuadrado y Correlaciones parcial y semiparcial.**
	Incluir múltiples variables da una explicación más completa de lo que está pasando. La valoración conjunta de múltiples variables permite eliminar del análisis efectos comunes o compartidos para poder captar el efecto neto de cada una de ellas. Incluir el modelo más **parsimonioso** con el **máximo ajuste.**	$Y = B_0 + B_1X + B_2X_2 \dots B_pX_p$ **Coeficientes de regresión** Son coeficientes de regresión parciales, pues el valor depende del resto de coeficientes incluidos en la ecuación.			
	VI Categóricas	**Regresión jerárquica o por pasos**	**Criterios para seleccionar variables**	**Supuestos**	
	Las variables dicotómicas se pueden incluir sin problema en la ecuación. Para las variables categóricas hay que crear para J-1 variables diferentes.	Se puede incluir en la ecuación todas las variables que se sospecha que pueden aportar algo, para luego ir desechando las que nos son significativas. Eliminar más de una variable al mismo tiempo impide valorar el comportamiento individual de las variables. Es por ello que se procede de forma **jerárquica**	1. Cuantificar el cambio que se produce en el coeficiente de determinación $\boldsymbol{R^2_{XY}}$. 2. Coeficientes en cada caso. 3. Variables excluidas **Métodos para seleccionar variables.** **Hacia delante:** se comienza con la intersección y se van añadiendo las variables con el coeficiente de correlación parcial más alto. **Hacia atrás:** se incluyen todas las variables candidatas y se procede a ir eliminándolas una a una. **Pasos sucesivos:** mezcla los dos modelos anteriores. 1. Se elige la variable que supera el criterio de selección y más alto correlaciona. 2. Cada vez que se incorpora una nueva variable se evalúa si hay alguna que no cumple el criterio de selección.	**Linealidad**: se asume que las variables independientes están linealmente relacionadas. **No colinealidad:** no debe haber relación directa entre dos VI. Indicios: 1. Cuando la F sea significativo y que no lo sea ninguno de los coeficientes de regresión. **Independencia:** Se mide con Durbin Watson 1,5-2,5 es adecuado. **Normalidad:** Cada valor de X le corresponde una población de valores de Y. Estas poblaciones son normales. **Homocedasticidad:**	