**GESTIÓN DEL CONOCIMIENTO EN LAS PYMES DEL VALLE DEL CAUCA**

**Metodología**

**Harold Achicanoy & César Saavedra**

A continuación se presenta una breve descripción de las técnicas estadísticas usadas para dar cumplimiento a los objetivos de análisis propuestos. Dada la naturaleza cualitativa de las variables se describen algunas técnicas específicas de análisis para datos categóricos, mientras también se presentan otras específicas para el análisis de datos numéricos que pueden ser adaptadas para datos categóricas.

La distribución de los temas se presenta de la siguiente manera: en la sección de **Análisis de correlación** se presentan las metodologías utilizadas en la cuantificación de relaciones existentes entre variables y también las relaciones existentes entre pilares, considerando estos como correlaciones entre tablas de datos. Por otro lado, se presentan las bases del modelo usado para el **Análisis de causalidad**, PLS-PM, el cual permite cuantificar el peso de las relaciones potenciales que pueden darse entre los pilares del conocimiento con base en un modelo teórico propuesto.

**Análisis de correlación**

***Test de independencia Chi-cuadrado***

El test de independencia Chi-cuadrado permite verificar si dos variables categóricas son independientes o no (es decir, si existe algún tipo de relación entre ellas lo que indicaría que una de ellas influye en la otra). Para esto se plantea el siguiente contraste estadístico de hipótesis:

vs

Para esto se realiza la construcción de una tabla de contingencia para el par de variables en comparación y se calculan las frecuencias observadas las cuales se contrastan con las frecuencias esperadas bajo la hipótesis de independencia. De tal manera, que si en dicho contraste las frecuencias observadas superan significativamente las esperadas bajo el supuesto de independencia se puede concluir que las variables bajo análisis no son independientes y presentan algún tipo de relación.

De este modo, para realizar la comparación entre todos los ítems del cuestionario dentro de cada pilar, se fijó un nivel de significancia de α=0.05 el cual fue contrastado con los valores-p generados de cada test específico. De tal modo que si el Valor-p < α indica que hay suficientes indicios que confirman algún tipo de relación entre el par de variables bajo análisis.

***Test V de Cramer***

Como complemento al test de independencia Chi-cuadrado surge el test V de Cramer el cual consta de una corrección del coeficiente Chi-cuadrado, permitiendo obtener un índice con valor máximo 1 (que indica la máxima asociación entre variables) y valor mínimo de 0 que indica que no existe ningún tipo de asociación, obteniendo un indicador de la fuerza de la asociación que existe entre las variables bajo análisis.

Este test se aplica posterior al test Chi-cuadrado para todos los pares de variables dentro de cada pilar en análisis.

***Análisis de correlación canónica***

El análisis de correlación canónica es una técnica estadística multivariada que explora las relaciones que puedan darse entre dos grupos de variables y su validez. Para el desarrollo del método, se crean combinaciones lineales de las variables originales, sobre la base de su estructura de correlación. A partir de este punto se aplica un procedimiento de descomposición matricial, similar al utilizado en el Análisis de Componentes Principales, pero buscando maximizar la correlación entre los dos grupos en estudio, de donde se pueden obtener los valores propios los cuales miden la correlación multivariada entre las nuevas variables lineales creadas a partir del análisis.

Para cada uno de los pares de pilares bajo consideración se realizaron los análisis de correlación canónica calculando el primer valor propio como indicador de la máxima correlación encontrada entre los pilares bajo estudio.

**Análisis de causalidad**

***Partial Least Squares – Path Modeling (PLS-PM)***

Contraparte no paramétrica de los modelos de ecuaciones estructurales (Structural Equation Models, SEM), flexible con los supuestos requeridos en los análisis paramétricos SEM. El objetivo principal de los modelos PLS-PM consiste en cuantificar la asociación a través de modelos de regresión entre variables latentes (o no observables) las cuales se miden a través de bloques de variables (Sanchez, 2013). Adicionalmente el modelo permite:

* Calcular un índice para cada bloque de variables bajo análisis, lo que permite cuantificar algún concepto clave no observable o variable latente de importancia que se desee medir.
* Examinar las relaciones entre variables latentes.
* Evaluar cuáles son las principales variables observadas que contribuyen a definir cada uno de los constructos o variables latentes.

El modelo PLS-PM consta de dos partes: el modelo interno (o estructural) y el modelo externo (o de medida). El modelo interno es la parte que explora las relaciones entre las variables latentes del modelo. Mientras, el modelo externo es el que cuantifica las relaciones entre cada variable latente con su propio bloque de variables observadas. En este punto se distinguen dos casos de especial interés: el primero denominado modelo reflectivo, donde la variable latente se considera una causa de las variables medidas. El otro caso corresponde al modelo formativo, en el cual las variables observadas se consideran como variables predictoras de la variable latente.

De este modo, la estimación conjunta del modelo se realiza mediante un procedimiento iterativo, en el cual de manera inicial se asignan aleatoriamente los pesos o coeficientes que permiten dar una cuantificación de las variables latentes y posteriormente se aplican modelos de regresión lineal múltiple, repitiendo este proceso hasta alcanzar un nivel de convergencia deseado, realizando una actualización de los coeficientes estimados en cada etapa.

La evaluación del modelo, inicia a través del modelo de medida o externo, en el cual dependiendo de su construcción (reflectivo o formativo), para el caso de un modelo formativo se examina el valor absoluto de las cargas estimadas y se prueba la posible existencia de multicolinealidad entre las variables medidas. Posteriormente, se continúa con la evaluación del modelo estructural donde se examinan los coeficientes de determinación y los coeficientes estimados mediante técnicas de remuestreo.

En este punto todos los modelos ajustados, se construyeron bajo el supuesto que los indicadores observados se generan a través de un modelo formativo.

**Bibliografia**

Sanchez, G. (2013). *PLS Path Modeling with R*.