

Keywords cluster analysis, factor analysis, factorial scoring, method of components principals, multivariate analysis.

Resumen El aporte principal del presente artículo, se basa en la tesis escrita por el Ing. Gandy Rene López Fuertes MSc., esta propuesta metodológica tiene como objetivo focalizar nuevos centros para la atención de personas con discapacidad, en este sentido se lleva a cabo la agrupación de las parroquias ecuatorianas con el fin de generar posibles escenarios para la creación de centros de atención para la población con discapacidad que demanden prioridad. La fuente principal para la elaboración de este modelo fue el VII Censo de población y VI censo de vivienda del año 2010, de dicha fuente se obtuvieron algunas variables sociodemográficas, como la incidencia de personas con discapacidad, personas con discapacidad que viven en pobreza por necesidades básicas insatisfechas, entre otras. Luego de aplicar métodos estadísticos multivariantes como el Análisis Factorial, se simplificaron algunas variables, y se tomó los dos primeros componentes, para lo cual fue posible explicar la varianza del 88,6% de los datos originales. Además, gracias a los resultados obtenidos del análisis de conglomerados, fue posible determinar los tres grupos de parroquias necesarios que mostraron cierto nivel de prioritaria. Finalmente, el grupo con más alta prioridad fue el Estrato 3, que corresponde a 378 (36,9%) parroquias cuyo conglomerado permitió clasificar a las parroquias más vulnerables en el Ecuador, en las que se necesita la creación de nuevos centros de atención, con la intención de que el Ministerio de Inclusión Económica y Social pueda asistir a las personas con discapacidad.

Palabras Claves análisis de conglomerados, análisis factorial, análisis multivariante, método de componentes principales, puntuaciones factoriales.

1. Introducción

El modelo social actual se basa en derechos y principalmente se centra en la persona con discapacidad y su entorno, el concepto de discapacidad es considerado como un problema asociado a las actitudes sociales y económicas de una sociedad, situación que impide la participación de la persona con discapacidad en la comunidad. En este sentido, el Estado debe garantizar el ejercicio de la política social, ejecutando políticas públicas para superar este problema, sin que se pierda de vista a los grupos prioritarios (CONADIS, 2013).

La política social, menciona que se debe disminuir la desigualdad entre los individuos, y así alcanzar la verdadera inclusión de los individuos. La propuesta de la política pública para este grupo prioritario, admite que la situación de discapacidad es una verdadera realidad, que posiblemente debe ser modificada transversalmente con la gestión interinstitucional, es así que el modelo actualmente se compone de procesos en los que se diferencian estas prácticas de inclusión (Asamblea Nacional, 2008).

En este contexto, según en el artículo 6 de la ley para este grupo vulnerable (Ley Orgánica de Discapacidades, 2012), menciona que un sujeto con discapacidad es como consecuencia de alguna deficiencia física, mental, entre otras e independientemente de la causa original que se encuentra proscrito de sus derechos.

Análogamente, en la Clasificación Internacional de Deficiencias, Discapacidades y Minusvalías (World Health Organization, 1994), se define a la discapacidad de una persona de forma específica, así, la discapacidad es la limitación o ausencia de capacidad para efectuar actividades según sus particularidades sociodemográficas debido a una enfermedad. En Sanchez (2010), explica que el término discapacidad se ha diferenciado en cualquier estrato poblacional, produciendo políticas significativas en la colectividad.

Por otra parte, la calidad de vida es conceptualizada como un estado deseado de bienestar personal, que se define a partir de los dominios e indicadores básicos que configuran el bienestar personal. En este aspecto, podemos obtener una concepción de la calidad de la vida, como la optimización de programas y servicios, pues este concepto se utiliza como un criterio para evaluar la efectividad de servicios (Giné, 2004).

López (2019) afirma que, para lograr la calidad de vida de las personas con discapacidad, es necesario focalizar y priorizar servicios públicos en parroquias en las cuales se necesita establecer nuevos centros de atención para personas con discapacidad.

Consecuentemente, al 2021, se deberá contar con infraestructura adecuada, para los servicios integrales de personas con discapacidad, esto implica una atención de calidad, como también prácticas de salud preventiva y articulación interinstitucional permanente.

Finalmente, el propósito de esta investigación es la de mostrar la propuesta metodológica que según López (2019), permite estratificar a las parroquias del Ecuador para la creación de nuevos centros de atención para personas con discapacidad del Ministerio de Inclusión Económica y Social a través de información censal, con el empleo de métodos estadísticos del análisis multivariante.

A fin de cuentas, para la aplicación de esta metodología se definieron variables socio-demográficas de la población con discapacidad a ser incorporadas en el modelo Factorial, que por medio de la técnica estadística del análisis multivariante Análisis Factorial se sintetizaron características de la población en estudio; transversalmente la técnica estadística del análisis multivariante denominada Análisis de Conglomerados, permitió formar las aglomeraciones de la población analizada de acuerdo a sus características.

El presente documento está organizado de la siguiente manera: Material y Métodos 2, Resultados 3, Discusión 4, Conclusiones 5, Trabajo Futuro 6.

2. Material y Métodos

En el modelo de estratificación de las parroquias del Ecuador, en el cual se propone las parroquias prioritarias y focalizadas donde deben ser construidos los centros para personas con discapacidades, se utilizó como fuente principal el VII Censo de Población y VI de Vivienda del año 2010 (INEC, 2010). Esta base de datos es completa representativa a nivel nacional, cuenta con 24 provincias, 224 cantones y 1.024 parroquias del Ecuador, entonces la unidad de observación es la parroquia, además, esta fuente permite obtener la información de las variables sociodemográficas del grupo vulnerable y en estudio a nivel de parroquia. Algo muy importante que se debe indicar, es que no se realizó ningún tipo de muestreo, por lo tanto, la población y la muestra son las 1.024 parroquias.

Así, según López (2019), se revisó el Cuestionario Censal de esta fuente de datos, de la cual se determinaron las variables cuantitativas sociodemográficas, como la incidencia de personas con discapacidad, incidencia de personas en pobreza por necesidades básicas insatisfechas, entre otras; para esto se identificaron algunas dimensiones y características, que dependen de la condición de vida y que tienen su efecto en la explicación de programas de prestación de servicios. En el Cuadro 1, se presentan los correspondientes indicadores de calidad de la Clasificación Internacional de Deficiencias, Discapacidades y Minusvalías (World Health Organization, 1994). Cabe recalcar que según López (2019) se indica la metodología para realizar el cálculo de los indicadores, y en la cual se diseñó el correspondiente código en el software estadístico IBM SPSS Statistics 25 (Visauta, 1998), (Vinacua, 1998) y Microsoft Power BI (Lachev y Price, 2018), (Ferrari y Russo, 2016) y (Powell, 2017).

Finalmente, se utilizaron técnicas estadísticas multivariantes como el Análisis Factorial el cual permitió reducir variables, y luego el Análisis de conglomerados, método Multivariante de ordenación automática de la información, el cual definió los grupos o estratos propuestos. En el modelo de estratificación de las parroquias del Ecuador, se elaboraron tres escenarios posibles de grupos o aglomeraciones de parroquias, en la cual se consideran prioritarias para la creación de nuevos centros para la atención de población con discapacidad.

En (Anderberg, 2014), (CATENA, 2003), (Cuadras, 2007), (Hair, Anderson, Tatham, y Black, 1999), (Kaiser, 1958), (Morales Vallejo, 2013), (Johnson y Wichern, 2002), (Uriel, 1995) se obtiene el marco conceptual de esta investigación, para lo cual comparten similares definiciones y enuncian que el análisis Multivariante es un conjunto de técnicas estadísticas que hacen un análisis simultáneo de dos o más variables.

Cuadro 1: Dimensiones e indicadores de calidad de vida considerados en la propuesta Metodológica, según Schalock y Verdugo

Dimensión	Variable	Descripción
Generales	X ₁	Número de personas con discapacidad.
Bienestar Físico	X ₂	Número de personas con discapacidad, de todas las edades, no afiliadas o cubiertas por los programas de seguridad social general del país en un año determinado.
	X ₃	Número de personas con discapacidad que se encuentran en pobreza por Necesidades Básicas Insatisfechas.
	X ₄	Número de hogares con personas con discapacidad que se encuentran en pobreza por Necesidades Básicas Insatisfechas.
	X ₅	Número de personas con discapacidad jefes de hogar que viven solas.
	X ₆	Porcentaje de la población con discapacidad de 15 años y más de edad que no puede leer, escribir, ni comprender un texto sencillo y corto sobre su vida cotidiana, en un periodo de tiempo.
Desarrollo Personal	X ₇	Número promedio de años escolares aprobados por la población con discapacidad de 24 años y más de edad en el sistema educativo formal en los niveles de educación general básica, educación bachillerato, educación superior universitaria o técnica/tecnológica y postgrado.
	X ₈	Número de personas con discapacidad, que no poseen ningún nivel de instrucción.
	X ₉	Número de personas con discapacidad que se encuentran cesantes o disponibles para trabajar.
	X ₁₀	Número de viviendas con personas con discapacidad, donde el régimen de propiedad de la vivienda es por servicios.
	X ₁₁	Número de viviendas con personas con discapacidad, donde el régimen de propiedad de la vivienda es arrendada o en anticresis.
Bienestar Material	X ₁₂	Número de viviendas con personas con discapacidad, donde la calidad de la vivienda es mala.
	X ₁₃	Número de viviendas con personas con discapacidad, donde la Obtención de agua entubada no es por red pública sino por otros medios.
	X ₁₄	Número de viviendas con personas con discapacidad, donde la vivienda no recibe agua por tubería sino por otros medios.
	X ₁₅	Número de viviendas con personas con discapacidad, donde la forma de eliminación de la basura es por otros medios.
	X ₁₆	Número de viviendas con personas con discapacidad, donde el servicio higiénico o escusado de la vivienda no tiene.

Fuente: (López, 2019)

3. Resultados

En López (2019), se presenta la sintaxis construida en el software estadístico SPSS relacionada a la determinación de los indicadores, además se indica de manera didáctica la forma de aplicar el Análisis Factorial y el Análisis de Conglomerados, finalmente su diseño y construcción en Microsoft Power Business Intelligence, líder a nivel mundial. A continuación, se presenta los resultados del Análisis Factorial y del Análisis de Conglomerados.

El propósito del Análisis Factorial, consiste en englobar el problema en una estructura matricial, de tal manera que se define la matriz de correlaciones ver la Figura 1 entre un gran número de variables, en Hair et al. (1999). Cuando el número de variables es demasiado grande, surge la necesidad de representar a través de un número de variables más pequeño. En consecuencia, el Análisis Factorial es el método indicado para realizar la selección de un grupo de variables mejor representativo, creando variables llamadas factores que remplazan a las variables originales.

Según Hair et al. (1999), el objetivo fundamental del Análisis Factorial es analizar la matriz de correlaciones, esta matriz debe determinar las relaciones posibles entre variables. Para que el Análisis Factorial sea muy ventajoso es necesario que en la matriz de correlaciones incluya variables que se encuentren fuertemente correlacionadas. Podemos concluir que una matriz de correlaciones lo más aproximada a la matriz identidad, implica que el Análisis Factorial converge a una solución deficiente.

Partimos por analizar la matriz de correlaciones, en la Figura 1, se muestra dicha matriz de correlaciones y los niveles de significancia, en el que se determinó que dicha matriz de correlaciones parciales a nivel bivariado, posee valores óptimamente mayores o iguales a 0,7 indicando una colinealidad alta, y su determinante es distinto de cero, e igual a $9,124 \times 10^{-26}$, lo cual es la condición inicial que debe cumplir el Análisis Factorial, implicando una multicolinealidad entre las variables.

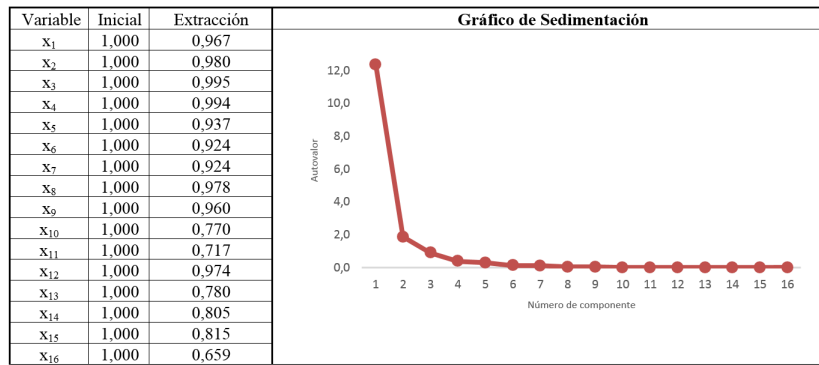
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16	
Correlación	X1	1.000	0.997	0.975	0.974	0.993	0.206	0.993	0.991	0.860	0.908	0.957	0.790	0.813	0.878	0.716	
	X2	0.997	1.000	0.984	0.984	0.984	0.120	0.203	0.991	0.994	0.848	0.877	0.972	0.813	0.840	0.904	0.738
	X3	0.975	0.984	1.000	1.000	0.956	0.118	0.203	0.982	0.975	0.850	0.815	0.992	0.895	0.908	0.906	0.797
	X4	0.974	0.984	1.000	1.000	0.954	0.118	0.204	0.980	0.975	0.847	0.810	0.993	0.896	0.911	0.909	0.798
	X5	0.993	0.984	0.956	0.954	1.000	0.115	0.202	0.987	0.975	0.868	0.938	0.930	0.760	0.768	0.831	0.698
	X6	0.123	0.120	0.118	0.118	0.115	1.000	0.868	0.095	0.118	0.182	0.131	0.111	0.082	0.092	0.079	-0.023
	X7	0.206	0.203	0.203	0.204	0.202	0.868	1.000	0.187	0.192	0.277	0.213	0.199	0.164	0.158	0.131	0.074
	X8	0.993	0.991	0.982	0.980	0.987	0.095	0.187	1.000	0.980	0.873	0.895	0.967	0.826	0.834	0.863	0.756
	X9	0.991	0.994	0.975	0.975	0.975	0.118	0.192	0.980	1.000	0.820	0.862	0.962	0.788	0.834	0.915	0.718
	X10	0.860	0.848	0.850	0.847	0.868	0.182	0.277	0.873	0.820	1.000	0.856	0.818	0.759	0.721	0.646	0.588
	X11	0.908	0.877	0.815	0.810	0.938	0.131	0.213	0.895	0.862	0.856	1.000	0.768	0.579	0.566	0.622	0.504
	X12	0.957	0.972	0.992	0.993	0.930	0.111	0.199	0.967	0.962	0.818	0.768	1.000	0.895	0.913	0.913	0.819
	X13	0.790	0.813	0.895	0.896	0.760	0.082	0.164	0.826	0.788	0.759	0.579	0.895	1.000	0.967	0.779	0.807
	X14	0.813	0.840	0.908	0.911	0.768	0.092	0.158	0.834	0.834	0.721	0.566	0.913	0.967	1.000	0.859	0.784
	X15	0.878	0.904	0.906	0.909	0.831	0.079	0.131	0.863	0.915	0.646	0.622	0.913	0.779	0.859	1.000	0.728
	X16	0.716	0.738	0.797	0.798	0.698	-0.023	0.074	0.756	0.718	0.588	0.504	0.819	0.807	0.784	0.728	1.000
Sig. (unilateral)	X1	0.000	0.000	0.000	0.000	0.000	0.077	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	X2	0.000	0.000	0.000	0.000	0.000	0.082	0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	X3	0.000	0.000	0.000	0.000	0.000	0.085	0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	X4	0.000	0.000	0.000	0.000	0.000	0.086	0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	X5	0.000	0.000	0.000	0.000	0.000	0.091	0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	X6	0.077	0.082	0.085	0.086	0.091	0.000	0.000	0.134	0.086	0.017	0.065	0.099	0.171	0.143	0.179	0.396
	X7	0.008	0.009	0.009	0.009	0.009	0.000	0.000	0.015	0.013	0.001	0.006	0.010	0.028	0.033	0.064	0.196
	X8	0.000	0.000	0.000	0.000	0.000	0.134	0.015	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	X9	0.000	0.000	0.000	0.000	0.000	0.086	0.013	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	X10	0.000	0.000	0.000	0.000	0.000	0.017	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	X11	0.000	0.000	0.000	0.000	0.000	0.065	0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	X12	0.000	0.000	0.000	0.000	0.000	0.099	0.010	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	X13	0.000	0.000	0.000	0.000	0.000	0.171	0.028	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	X14	0.000	0.000	0.000	0.000	0.000	0.143	0.033	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	X15	0.000	0.000	0.000	0.000	0.000	0.179	0.064	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	X16	0.000	0.000	0.000	0.000	0.000	0.396	0.196	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Figura 1: Matriz de Correlación y significancia
Fuente:(López, 2019)

El KMO (Kaiser-Meyer-Olkin) es un índice que compara cada valor de la matriz de correlación observado con el valor de la correlación parcial, así el estadístico KMO es 0,899 valor cercano a la unidad, lo que indica una adecuación excelente de los datos, y según el contraste de Bartlett refleja que no es significativa la hipótesis nula de variables iniciales incorreladas, con lo anterior se confirma la aplicación del Análisis Factorial. Por lo tanto, a pesar de que las variables (x_7) y (x_8) son no significativas, el KMO (Kaiser-Meyer-Olkin) y el contraste de Barlett confirman la aplicación del Análisis Factorial.

Según Hair et al. (1999), las comunalidades atribuidas a las variables iniciales y las comunalidades reproducidas por la solución factorial, indican que la comunalidad de un indicador es el porcentaje de su varianza que puede ser demostrada por el modelo factorial. Las comunalidades de la extracción se pueden determinar por los indicadores del modelo. Por lo tanto, según López (2019) la variable que mayormente explica es el número de personas con discapacidad en la parroquia (x_1) con el 96,7%, y el indicador que menos explica es el número de viviendas con personas con discapacidad, y finalmente el servicio higiénico o escusado de la vivienda no tienen (x_{16}) explica sólo el 65,9% de la variabilidad.

Además, según Hair et al. (1999), el método de extracción de componentes principales, es un método de extracción en el que supone la posibilidad de explicar el 100% de la varianza observada y, por ello, todas las comunalidades iniciales son idénticas a la unidad. Según López (2019), el método utilizado fue el de componentes principales, a partir de este se determinó que, las variables incluidas en el modelo son suficientes para realizar el análisis. En la Figura 2 se presentan las comunalidades y el correspondiente gráfico de sedimentación.



Método de extracción: análisis de componentes principales.

Figura 2: Comunalidades y Gráfico de Sedimentación
Fuente:(López, 2019)

Conforme a lo que menciona López (2019), del gráfico de sedimentación se obtiene el número de factores óptimo, es decir que el número óptimo de factores cuyos autovalores mayores a 1 se consideran como componentes o factores comunes, son

las dos primeras componentes. Esto significa que los autovalores mayores que 1, son las dos primeras componentes que logran explicar el 88,6% de la varianza de los datos.

De acuerdo con Anderberg (2014), el Análisis de Conglomerados Jerárquico agrupa casos o variables, de acuerdo al método de agrupación y a la medida de distancia. Este análisis empieza construyendo la matriz de distancias entre los casos o variables. El Análisis de Conglomerados Jerárquico es un procedimiento aglomerativo: El cual parte de los elementos individualmente considerados, luego va creando grupos hasta llegar a establecer un solo conglomerado, el cual está conformado por el total de elementos de la muestra.

En correspondencia a lo mencionado anteriormente por Anderberg (2014) y López (2019), comienza el análisis con todos los casos desunidos en conglomerados individuales, donde en la primera etapa se refiere a casos individuales, es así que, a partir de ese momento, los casos 737 y 788 estos dos casos conforman el aglomerado 737 y son indivisibles en las etapas siguientes. La columna denominada Coeficientes indica el valor de la distancia a la que se encuentran estos casos antes de la unión. En la etapa 2, la distancia de fusión entre los casos 744 y 880 vale aproximadamente 0 (0,063) lo que significa que se trata de casos con idénticas puntuaciones. Luego este procedimiento continua hasta formar los demás conglomerados. En la Figura 3, se presenta el Historial de conglomeración, en el cual se indica que el conglomerado combinado indica acerca de los conglomerados fusionados en cada etapa.

Etapa	Clúster combinado		Coeficientes	Primera aparición del clúster de etapa		Etapa siguiente
	Clúster 1	Clúster 2		Clúster 1	Clúster 2	
1	737	788	0,026	0	0	146
2	744	880	0,063	0	0	9
3	169	236	0,100	0	0	98
4	424	687	0,138	0	0	119
5	765	904	0,177	0	0	394
6	755	764	0,217	0	0	44
7	824	915	0,258	0	0	118
8	364	394	0,300	0	0	172
9	21	744	0,342	0	2	493
...
493	21	159	58,560	9	125	773

Figura 3: Historial de conglomeración (Aplicación)

Fuente:(López, 2019)

Analizando el esquema de conglomeración determinado por López (2019) para obtener los tres estratos. Se menciona que, en la primera etapa la parroquia San Isidro, del Cantón Morona, es el caso 737, y la parroquia Shimpis, cantón Logroño de la provincia Morona Santiago, de la zona de planificación 6, es el caso 788 se

fusionan. Una vez realizado el primer conglomerado, se vuelve a recalcular una nueva matriz de distancias con las 1.023 parroquias restantes, esto es las 1.023 parroquias y la agrupación (San Isidro - Shimpis). Este procedimiento lo realizamos recursivamente utilizando la misma metodología explicada por López (2019). En la Figura 4, se indica la Matriz de Distancias, esta permite determinar las distancias entre cada par de elementos. Las distancias se determinan de acuerdo con la medida de similitud o de disimilitud y depende además del método de conglomeración. A continuación, se muestra la matriz de distancias para los primeros casos, según la distancia euclidiana.

Caso	1:Case 1	2:Case 2	3:Case 3	4:Case 4	5:Case 5	6:Case 6	7:Case 7	8:Case 8	9:Case 9	10:Case 10	11:Case 11	12:Case 12	13:Case 13	14:Case 14	...	1024:Ca1024
1:Case 1	0,000	2,248	0,699	2,580	0,923	0,592	1,333	1,044	2,716	0,713	1,306	1,692	0,940	0,968	...	2,126
2:Case 2	2,248	0,000	2,386	3,363	2,608	2,445	2,601	2,981	3,308	2,637	2,967	2,605	2,816	2,752	...	3,680
3:Case 3	0,699	2,386	0,000	2,221	1,324	0,616	1,749	1,084	3,145	0,933	1,038	2,112	1,095	0,995	...	1,906
4:Case 4	2,580	3,363	2,221	0,000	3,192	2,394	3,593	2,551	4,894	2,862	2,553	3,882	2,855	2,567	...	2,306
5:Case 5	0,923	2,608	1,324	3,192	0,000	0,871	0,588	1,335	2,025	0,495	1,447	1,168	0,746	1,010	...	2,614
6:Case 6	0,592	2,445	0,616	2,394	0,871	0,000	1,356	1,032	2,760	0,505	0,995	1,823	0,712	0,713	...	2,107
7:Case 7	1,333	2,601	1,749	3,593	0,588	1,356	0,000	1,841	1,510	1,040	1,980	0,641	1,308	1,534	...	3,110
8:Case 8	1,044	2,981	1,084	2,551	1,335	1,032	1,841	0,000	3,316	1,046	0,882	2,256	0,790	0,583	...	1,295
9:Case 9	2,716	3,308	3,145	4,894	2,025	2,760	1,510	3,316	0,000	2,451	3,398	1,317	2,731	3,003	...	4,599
10:Case 10	0,713	2,637	0,933	2,862	0,495	0,505	1,040	1,046	2,451	0,000	1,043	1,576	0,470	0,728	...	2,278
11:Case 11	1,306	2,967	1,038	2,553	1,447	0,995	1,980	0,882	3,398	1,043	0,000	2,504	0,770	0,658	...	1,635
12:Case 12	1,692	2,605	2,112	3,882	1,168	1,823	0,641	2,256	1,317	1,576	2,504	0,000	1,862	2,028	...	3,473
13:Case 13	0,940	2,816	1,095	2,855	0,746	0,712	1,308	0,790	2,731	0,470	0,770	1,862	0,000	0,407	...	2,005
14:Case 14	0,968	2,752	0,995	2,567	1,010	0,713	1,534	0,583	3,003	0,728	0,658	2,028	0,407	0,000	...	1,711
...
1024:Ca1024	2,126	3,680	1,906	2,306	2,614	2,107	3,110	1,295	4,599	2,278	1,635	3,473	2,005	1,711	...	0,000

Figura 4: Matriz de Disimilitudes - Distancias entre los casos (Distancia Euclídea)
Fuente:(López, 2019)

Ahora se determina las medidas de proximidad entre los n casos por pares. El análisis se lo realiza para las 1.024 parroquias. Así, en la Figura 4 se muestran los coeficientes determinados utilizando la distancia euclidiana . Considerando la propuesta metodológica de López (2019), se utiliza la ecuación (1) de la distancia euclidiana para obtener el valor de la distancia, finalmente, en (6) se presenta el valor de la distancia 2,248; el cual representa la distancia euclídea entre el primer y segundo caso.

$$d(x_i, x_j) = \sqrt{\sum_{c=1}^p (x_{ic} - x_{jc})^2} \tag{1}$$

$$d(LAS\ GOLONDRINAS, MANGA\ DEL\ CURA) = \sqrt{\sum_{c=1}^{16} (x_{ic} - x_{jc})^2} \tag{2}$$

$$d(LAS\ GOLONDRINAS, MANGA\ DEL\ CURA) = \sqrt{(x_{i1} - x_{j1})^2 + \dots + (x_{i16} - x_{j16})^2} \tag{3}$$

$$= \sqrt{(Puntuacion\ Z(poblacionpcd2010)_{11} - Puntuacion\ Z(poblacionpcd2010)_{21})^2 + \dots + (Puntuacion\ Z(no\ tiene)_{1\ 16} - Puntuacion\ Z(no\ tiene)_{2\ 16})^2} \tag{4}$$

$$= \sqrt{(-0,112944721514304 - 0,104638142097195)^2 + \dots + (-0,509019272876407 - 0,112098649887252)^2} \tag{5}$$

$$= \sqrt{5,054123} = 2,248138 \approx 2,248 \tag{6}$$

En consecuencia, según lo mencionado por Hair et al. (1999) y Uriel (1995), el Análisis de Conglomerados Jerárquico, fundamentalmente el Método de Ward, selecciona grupos con mínima suma de cuadrados dentro de grupos. Por lo tanto, según lo mencionado por López (2019) de este procedimiento se tomaron 3 conglomerados, de tal forma que los elementos de cada estrato tienen características similares. En el Cuadro 2, se indica en resumen el ranking de las parroquias del Ecuador en las cuales se quiere focalizar para la creación de nuevos centros de atención para personas con discapacidad del Ministerio de Inclusión Económica y Social. Finalmente, considera que tres estratos es la solución satisfactoria de acuerdo a la aplicación de las técnicas multivariantes, además esta refleja la agrupación de los indicadores en las dos variables de clasificación. En la Figura 5 se presenta la distribución (ranking) de las parroquias del Ecuador.

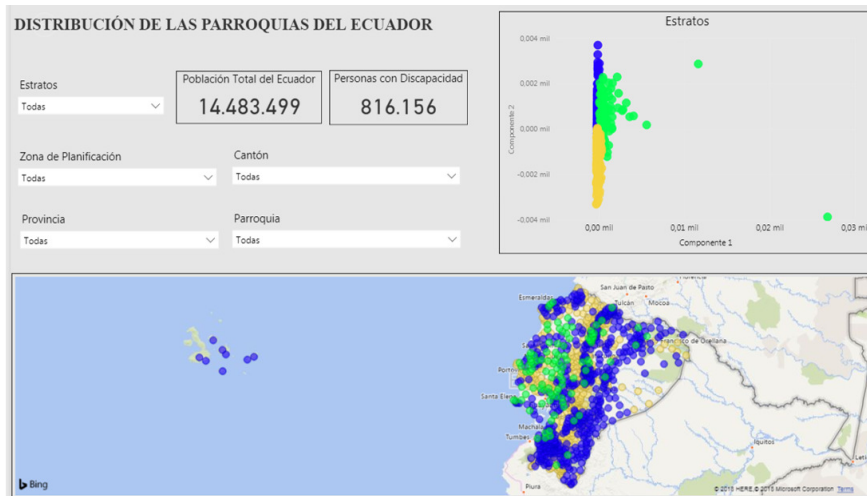


Figura 5: Distribución de las parroquias según el estrato
Fuente:(López, 2019)

Cuadro 2: Distribución del número de parroquias, según el estrato

Descripción	Estrato 1 (541)		Estrato 2 (105)		Estrato 3 (378)	
	Prioritario		No Prioritario		Muy Prioritario	
	μ (%)	σ	μ (%)	σ	μ (%)	σ
Porcentaje de personas con discapacidad.	6,5	2,0	5,9	1,1	7,2	2,1
Porcentaje de personas con discapacidad, de todas las edades, no afiliadas o cubiertas por los programas de seguridad social general del país en un año determinado.	64,0	11,6	67,1	7,3	67,8	11,9
Porcentaje de personas con discapacidad que se encuentran en pobreza por Necesidades Básicas Insatisfechas.	80,7	15,1	71,2	19,3	90,5	9,0
Porcentaje de hogares con personas con discapacidad que se encuentran en pobreza por Necesidades Básicas Insatisfechas.	16,7	5,5	13,5	5,1	19,7	4,8
Porcentaje de personas con discapacidad jefes de hogar que viven solas.	2,2	1,5	1,6	0,4	2,9	1,5
Porcentaje de la población con discapacidad de 15 años y más de edad que no puede leer, escribir, ni comprender un texto sencillo y corto sobre su vida cotidiana, en un periodo de tiempo.	74,9	6,5	75,2	8,0	55,7	9,7
Porcentaje promedio de años escolares aprobados por la población con discapacidad de 24 años y más de edad en el sistema educativo formal en los niveles de educación general básica, educación bachillerato, educación superior universitaria o técnica/tecnológica y postgrado.	5,1	0,9	5,8	1,3	3,2	0,7
Porcentaje de personas con discapacidad que no poseen ningún nivel de instrucción.	28,9	6,4	29,1	7,4	44,7	9,9
Porcentaje de personas con discapacidad que se encuentran cesantes o disponibles para trabajar.	1,0	1,1	1,6	0,7	0,6	1,0
Porcentaje de viviendas con personas con discapacidad, donde el régimen de propiedad de la vivienda es por servicios.	0,4	0,6	0,4	0,5	0,3	0,4
Porcentaje de viviendas con personas con discapacidad, donde el régimen de propiedad de la vivienda es arrendada o en anticresis.	1,4	1,1	2,3	1,3	0,7	0,7
Porcentaje de viviendas con personas con discapacidad, donde la calidad de la vivienda es mala.	7,4	3,5	5,7	2,6	9,6	3,9
Porcentaje de viviendas con personas con discapacidad, donde la Obtención de agua entubada no es por red pública sino por otros medios.	10,0	6,2	7,7	5,7	12,8	5,6
Porcentaje de viviendas con personas con discapacidad, donde la vivienda no recibe agua por tubería sino por otros medios.	4,7	5,2	4,8	4,2	5,6	5,1
Porcentaje de viviendas con personas con discapacidad, donde la forma de eliminación de la basura es por otros medios.	0,5	1,2	0,1	0,2	0,4	1,0
Porcentaje de viviendas con personas con discapacidad, donde el servicio higiénico o escusado de la vivienda no tiene.	4,2	3,7	1,6	1,4	6,4	4,4
Tasa de crecimiento	0,040	0,048	0,046	0,023	0,034	0,046

Fuente:(López, 2019)

4. Discusión

López (2019) indica que el total de habitantes en el país, según el VII censo de población, VI de vivienda del año 2010 INEC (2010), fue de 14.483.499; de los cuales 816.156 (5,6%) habitantes tenían algún tipo de discapacidad, 547.305 (67,1%) personas con discapacidad de todas las edades que no se encontraban afiliadas o cubiertas por los programas de seguridad social general, 520.567 (63,8%) personas con discapacidad que se encontraban en pobreza por Necesidades Básicas Insatisfechas, 694.831(85,1%) personas con discapacidad de 15 años y más de edad.

Además, conforme a lo mencionado por López (2019), existen 533.491 (65,4%) personas con discapacidad de 15 años y más de edad que no saben leer y escribir, lo que implica que la tasa de analfabetismo de este grupo prioritario es del 76,8%, además 614.548 (75,3%) personas con discapacidad tienen 24 años y más de edad. También, según el VII censo de población VI de Vivienda del año 2010 INEC (2010), el total de hogares del país fue de 3.810.548, de los cuales 661.319 (17,4%) hogares albergaban personas con discapacidad, 416.797(63%) hogares con personas con discapacidad, que se encontraban en pobreza por Necesidades Básicas Insatisfechas, 60.561(7,4%) hogares con personas con discapacidad jefes de hogar que viven solas. También, en consonancia con lo mencionado por López (2019) y según el INEC (2010), para el año 2010 se tiene que el total de viviendas en el país fue de 3.748.919, de las cuales 654.359 (17,5%) viviendas en los que habitan personas con discapacidad.

Ahora de acuerdo al CONADIS a marzo del 2020, existen 485.325 personas con discapacidad registradas, esto es el (59,5%) de personas con discapacidad a las cuales se conoce la situación sociodemográfica, como también su vulnerabilidad. Tomando en cuenta que dicho porcentaje solamente es subjetivo puesto que actualmente se desconoce la cantidad de personas con discapacidad, es decir existe un subregistro.

Los argumentos expuestos en el transcurso de este trabajo, revelan que efectivamente desde la condición sociodemográfica de la persona con discapacidad se encuentra limitada, no solo en su calidad de vida sino también en la atención en las instituciones públicas y/o privadas, puesto que, el presupuesto estatal para acciones destinadas a brindar accesibilidad universal para obtener la calidad de vida de las personas con discapacidad son escasos.

5. Conclusiones

En el presente estudio, se exploraron las principales dimensiones y características de las personas con discapacidad. En consecuencia, de las 91 variables originales se observó que el determinante era igual a 0, lo cual indica multicolinealidad en el conjunto de variables involucradas, para evitar esto se eliminó variables considerando criterios de privación o carencia en alguna de las dimensiones establecidas,

hasta obtener una matriz de 16 x 16 para lograr que el determinante de la matriz de correlaciones sea distinto de cero.

Al analizar la matriz de correlaciones, se determinó que dicha matriz de correlaciones parciales a nivel bivariado, posee valores óptimamente mayores o iguales a 0,7 indicando una alta colinealidad entre las variables, y su determinante es distinto de cero, e igual a $9,124 \times 10^{-26}$, lo cual implica multicolinealidad entre las variables, condición inicial que debe cumplir el análisis Factorial. De la matriz de correlaciones analizada entre las 16 variables e indicadores, se sintetizaron en 2 factores independientes. Los 2 factores extraídos logran explicar el 88,6% de la varianza total. A partir del Análisis de conglomerados, se construyeron 3 grupos de parroquias con cierto nivel de prioridad, así, el grupo muy prioritario (Estrato 3) corresponde a 378 (36,9%) parroquias, cuyo ordenamiento permitió estratificar a las parroquias del Ecuador para la creación de nuevos centros de atención para personas con discapacidad del Ministerio de Inclusión Económica y Social.

López (2019) manifiesta que, la tasa de crecimiento o decrecimiento obtenida para las personas con discapacidad, es también un indicador importante que debe incluirse en conjunto con las otras características, pues permite priorizar de mejor manera las parroquias del Ecuador. Puesto que, en cada grupo, hay parroquias donde existe una tasa alta de crecimiento de personas con discapacidad, tasa de crecimiento que permite generar un ranking dentro del conglomerado. Finalmente, el Estrato 3, es el estrato que posee mayor prioridad, pues los parámetros considerados en esta metodología son significativos en contraste a los otros dos estratos. En la provincia de Sucumbíos, Cantón Cascales y la parroquia Santa Rosa de Sucumbíos, es una de las parroquias prioritarias donde presenta una tasa de crecimiento intercensal para las personas con discapacidad del 20,24%.

6. Trabajo Futuro

En correspondencia, a lo mencionado por López (2019), en su propuesta metodológica Modelo Estratificado de las parroquias del Ecuador, en el que utiliza técnicas multivariantes, como el análisis factorial y el análisis de conglomerados para la creación de nuevos centros de atención para personas con discapacidad. En este sentido, tomando en cuenta la existencia de carencias de cuidado y atención de las personas con discapacidades, se tiene la necesidad de construir nuevos centros para la atención de personas con discapacidades, proyecto que es prioritario para el Estado; que a través de las instituciones agregadores de valor deben garantizar los derechos para grupos vulnerables, los cuales son amparados por la Constitución de la República y la ley Orgánica de Discapacidades.

Debido a la correlación entre discapacidad y pobreza no se ha alcanzado abordar de manera apropiada un nivel normativo entre estas dos características, el cual ha conllevado a un nivel de discriminación y exclusión social de las personas con discapacidad, para lo cual se propone utilizar esta metodología en acorde a lo que se establece en la política pública para las personas con discapacidad. En consecuen-

cia, a finales del año 2020 se realizará el VIII Censo de Población y VII de Vivienda, puesto que para el 2021 se podrá contar con información actualizada de las personas con discapacidad, y se pretenderá realizar la aplicación de esta metodología.

7. Bibliografía

Referencias

- Anderberg, M. R. (2014). *Cluster analysis for applications: probability and mathematical statistics: a series of monographs and textbooks* (Vol. 19). Academic press.
- Asamblea Nacional, A. C. d. E. (2008). Constitución de la República del Ecuador. *Quito: Tribunal Constitucional del Ecuador., 449.*
- CATENA, A. (2003). *Análisis multivariado: Un manual para investigadores. biblioteca nueva, España. 3]* chiappori, p. y gollier, c.(2006). *competitive failures in insurance markets: Theory and policy implications*. MIT Press, Estados Unidos.
- CONADIS. (2013). *Agenda Nacional para la Igualdad en Discapacidades 2013–2017*. Autor.
- Cuadras, C. M. (2007). *Nuevos métodos de análisis multivariante*. CMC editions Barcelona.
- Ferrari, A., y Russo, M. (2016). *Introducing microsoft power bi*. Microsoft Press.
- Giné, C. (2004). Servicios y calidad de vida para las personas con discapacidad intelectual. *Revista Española sobre Discapacidad Intelectual, 35(2)*, 1–14.
- Hair, J. F., Anderson, R. E., Tatham, R. L., e Black, W. C. (1999). *Análisis multivariante* (Vol. 491). Prentice Hall Madrid.
- INEC. (2010). Instituto nacional de estadísticas y censos. *Obtenido de INEC.*
- Johnson, R. A., y Wichern, D. W. (2002). *Applied multivariate statistical analysis* (Vol. 5) (n.º 8). Prentice hall Upper Saddle River, NJ.

- Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23(3), 187–200.
- Lachev, T., y Price, E. (2018). *Applied microsoft power bi: Bring your data to life!* Prologika Press.
- Ley Orgánica de Discapacidades, L. O. d. D. (2012). Ley Orgánica de Discapacidades. *Quito, Pichincha, Ecuador*.
- López, G., Fuertes. (2019). *Modelo estratificado de las parroquias del Ecuador mediante el análisis factorial y de conglomerados para la creación de nuevos centros de atención para personas con discapacidad* (Tesis de Master no publicada). Quito: UCE.
- Morales Vallejo, P. (2013). El análisis factorial en la construcción e interpretación de tests, escalas y cuestionarios. *Universidad Pontificia Comillas: Madrid*. Recuperado de: <http://web.upcomillas.es/personal/peter/investigacion/AnalisisFactorial.pdf>.
- Powell, B. (2017). *Microsoft power bi cookbook: Creating business intelligence solutions of analytical data models, reports, and dashboards*. Packt Publishing Ltd.
- Sanchez, G. (2010). Analisis y evaluacion de las politicas laborales y sociales en apoyo de los discapacitados en México. *Mexico: ELECTRONICA GRATUITA*.
- Uriel, E. J. (1995). *Análisis de datos: series temporales y análisis multivariante*.
- Vinacua, V. (1998). *Análisis estadístico con spss para windows. volumen ii*. McGRAW Hill, Madrid.
- Visauta, B. (1998). *Análisis estadístico con spss para windows. volumen i. estadística básica*. McGraw Hill: Madrid.
- World Health Organization, W. H. O. (1994). Clasificación internacional de deficiencias, discapacidades y minusvalías: manual de clasificación de las consecuencias de la enfermedad: publicada de acuerdo con la resolución wha29.35 de la vigesimonovena asamblea mundial de la salud, mayo 1976.