



CARTA DE AUTORIZACIÓN

CÓDIGO

AP-BIB-FO-06

VERSIÓN

1

VIGENCIA

2014

PÁGINA

1 de 1

Neiva, 23 de Julio de 2019

Señores

CENTRO DE INFORMACIÓN Y DOCUMENTACIÓN
UNIVERSIDAD SURCOLOMBIANA
Ciudad

Las suscritas Angie Paola Trujillo Montealegre, con C.C. No. 1075284220 y Livia Alejandra Caicedo Herrera, con C.C. No. 1082215886; autores de la tesis y/o trabajo de grado titulado **MODELO LINEAL GENERALIZADO DE LAS BRECHAS SALARIALES POR GÉNERO EN LA CIUDAD DE NEIVA: UNA APROXIMACIÓN CON REGRESIÓN USANDO ESTADÍSTICA BAYESIANA**, presentado y aprobado en el año 2019 como requisito para optar al título de Especialista en Estadística; autorizamos al CENTRO DE INFORMACIÓN Y DOCUMENTACIÓN de la Universidad Surcolombiana para que, con fines académicos, muestre al país y el exterior la producción intelectual de la Universidad Surcolombiana, a través de la visibilidad de su contenido de la siguiente manera:

- Los usuarios puedan consultar el contenido de este trabajo de grado en los sitios web que administra la Universidad, en bases de datos, repositorio digital, catálogos y en otros sitios web, redes y sistemas de información nacionales e internacionales "open access" y en las redes de información con las cuales tenga convenio la Institución.
- Permita la consulta, la reproducción y préstamo a los usuarios interesados en el contenido de este trabajo, para todos los usos que tengan finalidad académica, ya sea en formato Cd-Rom o digital desde internet, intranet, etc., y en general para cualquier formato conocido o por conocer, dentro de los términos establecidos en la Ley 23 de 1982, Ley 44 de 1993, Decisión Andina 351 de 1993, Decreto 460 de 1995 y demás normas generales sobre la materia.
- Continúo conservando los correspondientes derechos sin modificación o restricción alguna; puesto que, de acuerdo con la legislación colombiana aplicable, el presente es un acuerdo jurídico que en ningún caso conlleva la enajenación del derecho de autor y sus conexos.

De conformidad con lo establecido en el artículo 30 de la Ley 23 de 1982 y el artículo 11 de la Decisión Andina 351 de 1993, "Los derechos morales sobre el trabajo son propiedad de los autores", los cuales son irrenunciables, imprescriptibles, inembargables e inalienables.

ESTUDIANTE: **ANGIE PAOLA TRUJILLO MONTEALEGRE**

Firma: Angie Paola Trujillo H.

ESTUDIANTE: **LIVIA ALEJANDRA CAICEDO HERRERA**

Firma: Livia Alejandra Caicedo H.



TÍTULO COMPLETO DEL TRABAJO:

MODELO LINEAL GENERALIZADO DE LAS BRECHAS SALARIALES POR GENERO EN LA CIUDAD DE NEIVA: UNA APROXIMACIÓN CON REGRESIÓN USANDO ESTADISTICA BAYESIANA

AUTOR O AUTORES:

Primero y Segundo Apellido	Primero y Segundo Nombre
<ul style="list-style-type: none"> • Trujillo Montealegre • Caicedo Herrera 	<ul style="list-style-type: none"> • Angie Paola • Livia Alejandra

DIRECTOR Y CODIRECTOR TESIS:

Primero y Segundo Apellido	Primero y Segundo Nombre
Cangrejo Esquivel	Álvaro Javier

ASESOR (ES):

Primero y Segundo Apellido	Primero y Segundo Nombre
Cangrejo Esquivel	Álvaro Javier

PARA OPTAR AL TÍTULO DE: Especialista en Estadística.

FACULTAD: Ciencias Exactas y Naturales.

PROGRAMA O POSGRADO: Especialización en Estadística.

CIUDAD: Neiva

AÑO DE PRESENTACIÓN: 2019

NÚMERO DE PÁGINAS: 30

TIPO DE ILUSTRACIONES (Marcar con una X):

Diagramas Fotografías Grabaciones en discos Ilustraciones en general Grabados
Láminas Litografías Mapas Música impresa Planos Retratos Sin ilustraciones Tablas
o Cuadros



SOFTWARE requerido y/o especializado para la lectura del documento:

MATERIAL ANEXO:

PREMIO O DISTINCIÓN (En caso de ser LAUREADAS o Meritoria):

PALABRAS CLAVES EN ESPAÑOL E INGLÉS:

Español

Inglés

- | | |
|---------------------------|------------------------|
| 1. GLM | GLM |
| 2. Inferencia Bayesiana | Bayesian inference |
| 3. Distribución Gamma | Gamma distribution |
| 4. Distribución Gaussiana | Gaussiana distribution |

RESUMEN DEL CONTENIDO: (Máximo 250 palabras)

La presente investigación surge a partir de la discriminación salarial que se viene presentando entre hombres y mujeres en la ciudad de Neiva, haciendo uso de los datos obtenidos de la Misión para el Empalme de las Series de Empleo, Pobreza y Desigualdad (Mesep) 2016, con el fin de abarcar los propósitos de los individuos aplicando el método lineal generalizado por medio de inferencia bayesiana, encontrando que las ocupaciones por discriminación es uno de los factores que determinan lo que hoy llamamos discriminación salarial. De otro lado la morosa incorporación de la mujer al mercado forma el origen esencial de las diferencias salariales por género, así el estudio de las reformas laborales por genero presenta un gran interés en nuestra sociedad, donde notamos que el trabajo a lo largo de la historia ha ido mutando de manera signicativa en relación a la dependencia trabajador/capitalista, y en cierto modo, quienes realizan actividad laboral han ido conquistando a lo largo del siglo XX, diferentes derechos que les corresponden por su condición de asalariados, de igual forma es uno de los ejes fundamentales tanto de la vida social como la vida privada de un individuo.



ABSTRACT: (Máximo 250 palabras)

This research arises from the wage discrimination that has been presented between men and women in the city of Neiva, using the data obtained from the Mission for the Splice of the Employment, Poverty and Inequality Series (Mesep) 2016, with the to encompass the purposes of individuals by applying the generalized linear method through Bayesian inference, finding that occupations by discrimination is one of the factors that determine what we now call wage discrimination. On the other hand, the slow going of women into the market forms the essential origin of gender pay gaps, so the study of gender labour reforms is of great interest in our society, where we notice that work throughout the history has been mutating significantly in relation to worker/capitalist dependence, and in a way, those who engage in labor activity have been conquering throughout the twentieth century, different rights that correspond to them because of their status as wage earners, likewise it is one of the fundamental axes of both social life and the private life of an individual.

APROBACION DE LA TESIS

Nombre Presidente Jurado: Jaime Polanía Perdomo

Firma:

Nombre Jurado: Carlos Arturo Monje Alvarez

Firma:

**MODELO LINEAL GENERALIZADO DE LAS BRECHAS
SALARIALES POR GENERO EN LA CIUDAD DE NEIVA:
UNA APROXIMACIÓN CON REGRESIÓN USANDO
ESTADÍSTICA BAYESIANA.**

Presentado por:

**Angie Paola Trujillo Montealegre
Codigo 20182174815**

**Livia Alejandra Caicedo Herrera
Codigo 20182174813**

Tutor:

Alvaro Javier Cangrejo Esquivel




**Universidad Surcolombiana
Facultad de Ciencias Exactas y Naturales
Especialización en Estadística
Neiva - Huila
2019**

**FACULTAD DE CIENCIAS EXACTAS Y NATURALES
ESPECIALIZACIÓN EN ESTADÍSTICA**

CARTA DE ACEPTACIÓN

En calidad de Coordinador del Posgrado Especialización en Estadística, programa reconocido por el Ministerio de Educación Nacional mediante Resolución de Registro Calificado No. 3683 del 2 de marzo de 2018 y adscrito a la Facultad de Ciencias Exactas y Naturales de la Universidad Surcolombiana, me permito informar que el trabajo de investigación titulado: **“MODELO LINEAL GENERALIZADO DE LAS BRECHAS SALARIALES POR GÉNERO EN LA CIUDAD DE NEIVA: UNA APROXIMACIÓN CON REGRESIÓN USANDO ESTADÍSTICA BAYESIANA”** presentado por las estudiantes Angie Paola Trujillo Montealegre y Livia Alejandra Caicedo Herrera; es **ACEPTADO** como trabajo de grado para optar el título de Especialista en Estadística.

Para constancia se firma en la Ciudad de Neiva, a los diecinueve (19) días del mes de julio del año 2019.


JAIME POLANIA PERDOMO
Coordinador

CONSTRUYAMOS UNIVERSIDAD PARA EL DESARROLLO Y EL BUEN VIVIR

📍 Sede Central / Av. Pastrana Borrero - Cra. 1 ☎ PBX: 875 4753
📍 Sede Administrativa / Cra. 5 No. 23 - 40 ☎ PBX: 875 3686
🌐 www.usco.edu.co / Neiva - Huila ☎ Línea Gratuita Nacional: 018000 968722



AGRADECIMIENTOS

Le agradecemos a Dios por habernos acompañado y guiado a lo largo de la carrera, por ser nuestra fortaleza en los momentos de debilidad y por brindarnos una vida llena de aprendizajes, experiencias y sobre todo felicidad.

Le damos gracias a nuestros padres por apoyarnos en todo momento, por los valores que nos han inculcado, y por habernos dado la oportunidad de tener una educación en el transcurso de nuestra vida. Sobre todo por ser un ejemplo de vida a seguir.

RESUMEN

La presente investigación surge a partir de la discriminación salarial que se viene presentando entre hombres y mujeres en la ciudad de Neiva, haciendo uso de los datos obtenidos de la Misión para el Empalme de las Series de Empleo, Pobreza y Desigualdad (Mesep) 2016, con el fin de abarcar los propósitos de los individuos aplicando el método lineal generalizado por medio de inferencia bayesiana, encontrando que las ocupaciones por discriminación es uno de los factores que determinan lo que hoy llamamos discriminación salarial. De otro lado la morosa incorporación de la mujer al mercado forma el origen esencial de las diferencias salariales por género, así el estudio de las reformas laborales por género presenta un gran interés en nuestra sociedad, donde notamos que el trabajo a lo largo de la historia ha ido mutando de manera significativa en relación a la dependencia trabajador/capitalista, y en cierto modo, quienes realizan actividad laboral han ido conquistando a lo largo del siglo XX, diferentes derechos que les corresponden por su condición de asalariados, de igual forma es uno de los ejes fundamentales tanto de la vida social como la vida privada de un individuo.

Palabras claves: GLM, inferencia bayesiana, distribución gamma, distribución gaussiana.

ABSTRACT

This research arises from the wage discrimination that has been presented between men and women in the city of Neiva, using the data obtained from the Mission for the Splice of the Employment, Poverty and Inequality Series (Mesep) 2016, with the to encompass the purposes of individuals by applying the generalized linear method through Bayesian inference, finding that occupations by discrimination is one of the factors that determine what we now call wage discrimination. On the other hand, the slow going of women into the market forms the essential origin of gender pay gaps, so the study of gender labour reforms is of great interest in our society, where we notice that work throughout the history has been mutating significantly in relation to worker/capitalist dependence, and in a way, those who engage in labor activity have been conquering throughout the twentieth century, different rights that correspond to them because of their status as wage earners, likewise it is one of the fundamental axes of both social life and the private life of an individual.

Índice

1. EL PROBLEMA	9
1.1. JUSTIFICACIÓN	9
1.2. PLANTEAMIENTO DEL PROBLEMA	10
1.3. FORMULACIÓN DEL PROBLEMA	11
1.4. OBJETIVOS	12
1.4.1. Objetivo general	12
1.4.2. Objetivos específicos:	12
2. REFERENTES TEORICOS	13
2.1. Antecedentes de la Investigación	13
2.1.1. Internacional	13
2.1.2. Nacional	14
2.1.3. Regional	14
2.2. Marco Teorico	15
3. METODOLOGIA	19
3.1. Tipo de Investigación	19
3.2. Población	19
3.3. Diseño Muestral	19
3.4. Instrumentos	19
3.5. Procesamiento y análisis de información	19
4. ANÁLISIS DE RESULTADOS	25
5. CONCLUSIONES	29
6. RECOMENDACIONES	30
7. BIBLIOGRAFIA	31

INTRODUCCIÓN

La discriminación es un problema muy antiguo, pues desde el comienzo de la historia los hombres se han empeñado en crear separaciones entre ellos. Desde la servidumbre hasta la discriminación por razas y sexo; el ser humano lleva luchando con este concepto desde el inicio de su existencia e incluso hoy, en la era postmoderna, los conflictos discriminatorios continúan.

Colombia a través de los años ha reflejado una tasa de desempleo bastante significativa, su ascenso se debe en gran parte a su inestabilidad económica, política y social; actualmente en el campo laboral se produce la "discriminación laboral". Esto sucede cuando un individuo es criticado, menospreciado, vejado o maltratado física o emocionalmente; debido a sus creencias religiosas, costumbres, cultura, raza o nacionalidad, sexo u orientación sexual, e incluso la edad. Por tal motivo, se han creado leyes en cada país para defender a sus ciudadanos y evitar que sean discriminados, un factor importante para tener en cuenta es el desconocimiento sobre las reglas en materia laboral por parte de los trabajadores, ya que cierto conocimiento de dichas normas depende de que se respeten los derechos del trabajador y lograr una relación conforme entre empleado y empleador.

De otro lado, la afirmación de que el problema de informalidad laboral afecta más a la población femenina, se soporta en estudios como el liderado por el "PLAN INTEGRAL DECENAL DE IGUALDAD DE OPORTUNIDADES" donde el Gobierno Nacional y Regional ha motivado a adoptar políticas públicas para avanzar hacia la equidad e igualdad a través de programas, estrategias y acciones concretas que contribuyan a crear condiciones para terminar con esa desigualdad y permitan relaciones más democráticas en la vida pública y privada. La política Pública de Equidad de Género para las Mujeres Huilenses adoptada, mediante Ordenanza 013 de mayo 30 de 2014 permite impulsar acciones que garanticen el desarrollo, al reconocerle a las Mujeres un papel clave como agentes de este, reconoce sus derechos, su protagonismo social y político como ciudadanas y establece claramente la necesidad de formular el Plan Integral decenal de Igualdad de oportunidades. Así mismo el Plan Decenal es el instrumento a través del cual el Gobierno Departamental, define en aquellos ámbitos que son de su competencia, los objetivos y medidas prioritarias para eliminar cualquier discriminación por razón de sexo que pueda persistir y alcanzar la igualdad de oportunidades entre mujeres y hombres.

Así mismo el movimiento de las mujeres huilenses es un fenómeno reciente producto de la Constitución de 1991, principalmente. De manera análoga, es necesario entender que, desde los tiempos precolombinos, ya las mujeres se debatían en la lucha por la igualdad. Por consiguiente, cabe hacer memoria de la Cacica Gaitana, aquel personaje de resistencia indígena contra los invasores europeos, manifestación de ira, de dolor y de deshonra que puede sentir solo una madre por un hijo. Aquella prueba de la gran fuerza del pueblo latinoamericano enfrentándose contra los colonizadores que al descubrir el mundo nuevo lleno de campos verdes, oro, esmeraldas, cacao, tabaco, frutas exóticas, fauna y flora, querían arrasarlo en nombre de una civilización bárbara y cruel, es quien muestra la valentía y ardentía con que hacía valer sus derechos, no sólo frente al invasor español, a quien derrotó en gesta memorable, sino también frente a los demás caciques del sur del Huila.

De otro lado uno de los problemas laborales mayores de la mujer Huilense es la informalidad laboral que corresponde a diferentes modalidades:

- a) Venta ambulante y estacionaria informal que es el recurso más común de una economía de subsistencia para las personas que carecen de empresas propias o empleos formales para subsistir.
- b) Labores en empresas que no se encuentran debidamente legalizadas.
- c) Labores por la modalidad de peonazgo, o "trabajo por días" que resulta la forma más común de vinculación laboral

en el campo, en el sector minero y en el servicio doméstico.

Son tan graves estas situaciones laborales, que, según estudios realizados por el SENA, afectan a más del 50 % de la población económicamente activa. Con el agravante que, según otros estudios, la mayoría de esta población corresponde al sector femenino. Baste tener en cuenta la siguiente información publicada recientemente por el periódico La Nación, en referencia a la ciudad de Neiva, tendencia que puede extrapolarse a todo el departamento del Huila:

”Neiva cuenta con una tasa de desempleo que supera del 12 por ciento, según el reporte oficial del DANE. Sin embargo, el estudio realizado por el SENA indica que los índices de informalidad podrían superar el 25 por ciento. Y si a esta cifra se le suma el número de personas que desarrolla su actividad productiva dentro de parámetros legales, la informalidad laboral podría superar un indicador de 50 por ciento, evidenciando que los neivanos mayoritariamente viven de la informalidad laboral. Informalidad no son sólo ambulantes. El estudio destaca que la economía informal no está compuesta únicamente por los vendedores ambulantes, sino por el grupo de trabajadores independientes, empresas unipersonales o sociedades de hecho que operan normalmente por fuera de los parámetros del régimen fiscal impositivo, con alta inestabilidad y rotación laboral por fuera del sistema de seguridad social, sin cumplir con la normatividad del registro mercantil, normalmente sin acceder a los servicios del sistema financiero y con una inexistente práctica contable. ”Desde que no se cumpla con los parámetros establecidos por la Organización Internacional del Trabajo (OIT), es decir, con contratos serios con todas las prestaciones sociales, que tengan seguridad social, se considera que es un empleado que está por fuera de los términos legales y se puede llamar informal. En Neiva un tema que se venía dando con frecuencia era la contratación de personal por medio de cooperativas de trabajo asociado, con compensación cooperativa e incluso desconociendo otras compensaciones”, explicó. Informalidad empresarial, asimismo, el coordinador regional de formación profesional y empleo del SENA destacó que muchas empresas operan de forma informal, especialmente en el sector de la minería y en el área agropecuaria, imposibilitando que les brinden las garantías de ley a sus trabajadores. Un sector que goza de mucha informalidad es la minería, al igual que el sector agropecuario, pues a la persona se le paga un jornal sin ninguna contratación laboral, incluso muchas empresas agropecuarias no están formalizadas, entonces, aunque son empresas que operan en fincas contribuyen al sector empresarial del campo, perjudicando la formalización del empleo y de la empresa en el Huila”, puntualizó miguel Alirio Argote.

De lo anterior se evidencia un gran interés de estudio debido a que favorece la superación de pobreza, inclusión social y el desarrollo del país y la región; este trabajo desea conocer el porcentaje de empleo por género de los individuos a partir de un índice sistémico a través de las siguientes dimensiones: género, nivel de estudio, ingreso total.

El propósito de esta investigación es la de estimar el parámetro aplicando inferencia bayesiana para determinar la diferencia salarial en la ciudad de Neiva durante el año 2016 haciendo uso de los datos obtenidos de la Misión para el Empalme de las Series de Empleo, Pobreza y Desigualdad (Mesep) 2016, generando una serie de resultados utilizando el método lineal generalizado.

El presente trabajo está dividido en ocho apartados, de los cuales el primero se centra en la introducción, seguido del planteamiento del problema, el tercero se orienta al referente teórico, en el cuarto encontraremos la metodología, seguido encontraremos los resultados, contunamante en la sexta parte las conclusiones generales, posteriormente estan las recomendaciones y por último la bibliografía.

1. EL PROBLEMA

Desde los años 80's Colombia se ha destacado por tener altas tasas de participación femenina comparadas con otros países de América Latina y esa tendencia se ha mantenido hasta el presente. En los últimos 10 años las participaciones femeninas aumentaron del 54,4 % en 2008 a 59,1 % en 2016. Por su parte la participación masculina se ha mantenido más o menos estable alrededor de 75 %. Como resultado de estos cambios hubo una disminución en la brecha de participación de 18 a 15 puntos porcentuales. Fundamentalmente, las mujeres cada vez ganan más terreno en el mercado laboral, lo que significa la importancia que está tomando la mujer en cuanto a la participación en el mercado laboral.

La discriminación salarial por género constituye un ejemplo ilustrativo. Aunque las regulaciones jurídicas prevén la igualdad de remuneraciones por género, se siguen evidenciando las diferencias salariales al momento de remunerar el trabajo del hombre y la mujer. Por ejemplo, la tasa de paro de las mujeres es más elevada que la de los hombres, y las mujeres están subrepresentadas en empleos menos valorizados, como el trabajo temporal y el trabajo a tiempo parcial. Los factores que explican estas razones son generalmente: -El cuidado de los niños que normalmente recae mucho más sobre las mujeres que en los hombres y que constituye un obstáculo para el desarrollo de su carrera profesional. Prejuicios sobre los niveles de productividad de las mujeres trabajadoras en las empresas. De otro lado, los salarios que reciben las mujeres son por debajo que al de los hombres aproximadamente en el 20 %. Una gran parte de esta diferencia se explica por las características de los puestos de trabajo ocupados por las mujeres (tiempo parcial), su mayor presencia en empleos menos cualificados o trabajando en sectores con débil remuneración junto con interrupciones de su carrera profesional mucho más largas y numerosas.

Las mujeres siempre se han destacado por ocupar un puesto y ser independientes, efectuando cada vez más actividades remuneradas. En este sentido, el aumento de la tasa de actividad de las mujeres es una de las grandes transformaciones sociales de la segunda mitad del siglo XX.

1.1. JUSTIFICACIÓN

Las decisiones de política económica en materia de generación de empleo para las mujeres están relacionadas con la intervención que deben hacer los gobiernos desde el nivel local, departamental y nacional en el mercado de trabajo para disminuir el desempleo de las mujeres y mejorar las condiciones de permanencia y calidad de quienes ya tienen uno. Las políticas de empleo están necesariamente acompañadas de una movilización de recursos económicos, legales, técnicos y humanos para la obtención de unos resultados concretos, y generalmente implica no sólo la esfera pública si no también la esfera privada.

El modelo profesional de distribución de la población femenina en nuestro país está aún muy focalizado en ciertas actividades, especialmente, las que tienen que ver con la atención a los demás (educación, sanidad, cuidado de enfermos, ancianos), el comercio y los servicios; la tasa de paro femenina sigue manteniéndose por encima de la masculina, y todavía hay fuertes obstáculos a la plena incorporación de la mujer al trabajo (Drakich y Stewart, 2007).

Según datos del Dane para el 2017, por cada 100 hombres trabajadores asalariados existían 80 mujeres laborando en estas condiciones, de tal forma que ellas están mejor educadas, pero menos remuneradas mensualmente que los hombres. Por tanto las mujeres asalariadas tienen un año más de educación (12,16 años en promedio) que los hombres en sus mismas condiciones (10,91 años). Sin embargo, esto no ocurre con los emprendedores e independientes, donde los hombres tienen poco menos de un año de formación adicional que las mujeres, por estas razones nuestro

estudio se enfoca en determinar las probabilidades de empleo que tiene el género femenino en la ciudad de Neiva para el año 2016.

1.2. PLANTEAMIENTO DEL PROBLEMA

Las desigualdades del mercado laboral se refieren a las diferentes condiciones de vida que se dan entre individuos u hogares según una variable de referencia (renta o nivel de vida), por ejemplo, distribución de la riqueza mientras que las discriminaciones corresponden a una diferencia de trato entre dos personas sobre la base de un criterio no legal. Así, hay discriminación sobre el mercado laboral cuando dos personas con características productivas idénticas tienen un salario o una proposición de trabajo diferente como consecuencia de características no productivas diferentes, como es el caso, de la edad, sexo, nacionalidad, origen étnico-racial, entre otras.

Las personas con rentas más bajas son las más expuestas al riesgo de encontrarse en situación de desempleo de larga duración, mientras que las rentas medias o elevadas son las que normalmente han realizado estudios universitarios sin pasar por periodos de paro prolongados (Anne y L'Horty, 2013).

De acuerdo con (Braquet, 2014) quien afirma que el análisis económico muestra que los comportamientos discriminatorios plantean problemas jurídicos (son ilegales), éticos (crean desigualdades de trato injustificadas) y económicos (los trabajadores no son remunerados en función de su productividad marginal y de su capital humano). Además, las discriminaciones son complejas tienen múltiples facetas y están profundamente arraigadas en la conducta.

Según la OIT lo primordial es promover oportunidades para que mujeres y hombres obtengan un trabajo decente en condiciones de libertad, equidad, seguridad y dignidad humana. La visión de la OIT sobre la igualdad de género considera que se trata de un derecho humano fundamental, esencial para alcanzar el objetivo mundial de trabajo decente para todos.

La Política de la OIT sobre igualdad de género e incorporación de las consideraciones de género promueve dos líneas de acción: analizar y tratar en forma sistemática las necesidades específicas de hombres y mujeres, y realizar intervenciones dirigidas a grupos específicos con el fin de lograr la participación tanto de mujeres como de hombres, y para conseguir que se beneficien por igual de los esfuerzos por alcanzar el desarrollo.

Cientos de millones de personas son víctimas de la discriminación en el mundo del trabajo. Esto viola derechos humanos fundamentales, y además tiene profundas consecuencias económicas y sociales. La discriminación sofoca las oportunidades, desperdicia un talento humano que es necesario para el progreso, y acentúa las tensiones y desigualdades sociales. Combatir la discriminación es un componente esencial del trabajo decente, y los logros repercuten mucho más allá del lugar de trabajo.

Los temas relacionados con la discriminación están presentes en todo el trabajo de la OIT. Al promover la libertad sindical, por ejemplo, la OIT busca prevenir la discriminación contra los sindicatos y sus dirigentes. Los programas para combatir el trabajo forzoso y el trabajo infantil incluyen el apoyo a niñas y mujeres que están atrapadas en redes de prostitución o explotadas en servicio doméstico obligatorio. La no discriminación es el principio más importante del repertorio de recomendaciones prácticas sobre VIH/SIDA y el mundo del trabajo. Las directrices de la OIT sobre legislación laboral incluyen el tema de la discriminación, y en países como Namibia y Sudáfrica se ha realizado asesoría sobre cambios legislativos en esta área.

1.3. FORMULACIÓN DEL PROBLEMA

Los sentimientos de un empleado acerca de su sueldo y beneficios extrasalariales están influenciados por una variedad de factores que van más allá de la cantidad recibida y la calidad de los beneficios. Las comparaciones con otros, las expectativas, la propia deuda, las expectativas de los cónyuges, los niños y los compañeros, todo impacta en lo que sentimos acerca de cuánto se nos paga. No obstante, a pesar de eso, en la actualidad existen serios problemas en materia de remuneración laboral. Así, siguen existiendo desigualdades respecto a hombres y mujeres. Y es que los estudios encontrados afirman que aun estando en el mismo puesto y realizando los mismos trabajos, el hombre es mejor remunerado que la mujer.

Situación que demuestra que, aunque se haya luchado mucho y se hayan ido dando pasos importantes, se siguen observando desigualdades y discriminación.

Por tanto, surge la necesidad de poder reducir las brechas salariales entre hombres y mujeres, y los resultados que se presenten puedan ser visto y aplicados a aquellas empresas generadoras de empleo en la ciudad de Neiva, de igual forma generar gran interés para los entes gubernamentales de la ciudad de Neiva.

Siendo así podemos manifestar que a pesar del aumento participativo que ha tenido el género femenino, se sigue teniendo una gran diferencia en el campo laboral con respecto al género masculino; de lo anterior se indaga:

¿Cuál es el mejor modelo que se ajusta a los parametros de la brecha salarial por genero en la ciudad de Neiva obtenidas de la Misión para el Empalme de las Series de Empleo, Pobreza y Desigualdad (Mesep) 2016?.

1.4. OBJETIVOS

1.4.1. Objetivo general

Estimar el parámetro que permita establecer las principales determinantes de empleo por género en la ciudad de Neiva, usando métodos propios de la estadística bayesiana.

1.4.2. Objetivos específicos:

- Modelar el comportamiento de los ingresos salariales por género, asumiendo una distribución gaussiana.
- Modelar el comportamiento de los ingresos salariales por género, asumiendo una distribución gamma.
- Comparar los resultados obtenidos bajo las diferentes aproximaciones.

2. REFERENTES TEORICOS

2.1. Antecedentes de la Investigación

A través del tiempo se evidencia la desigualdad o la brecha salarial la cual no existe solo en Colombia si no a nivel mundial; existen varios factores que influyen en la brecha salarial por género, algunos de los factores son su característica del individuo en términos de productividad, los ambientes laborales y la cultura de las regiones. Cuando dos individuos (hombre y mujer), presentan las mismas características y habilidades, y aun estando en el mismo trabajo remuneran ingresos diferentes. Según (Stiglitz, 1973), las diferencias están sistemáticamente correlacionadas con ciertas características no económicas de dichas personas (raza, sexo, religión). A continuación, se observarán algunas investigaciones a nivel internacional, nacional y regional.

2.1.1. Internacional

A nivel internacional nos encontramos con la investigación que lleva por nombre "DISCRIMINACION SALARIAL POR SEXO: UN ANALISIS DEL SECTOR PRIVADO Y SUS DIFERENCIAS REGIONALES EN ESPAÑA" del cual son autores "Ricardo Alaez Aller - Miren Ullibarri Arce" el cual fue publicado en el año 2001; dicha investigación nos ofrece una medida tanto de las diferencias como de la discriminación salarial por sexo en las comunidades autónomas. Utilizando los datos de la Encuesta de Estructura Salarial, referidos al sector privado en el año 1995, las diferencias salariales se descomponen siguiendo el método de Oaxaca. Los resultados reflejan que, a nivel regional, unas mayores diferencias salariales no se corresponden necesariamente con una discriminación más elevada. Por otro lado, la incorporación tardía de la mujer al mercado laboral, así como su segregación sectorial y ocupacional, constituyen el origen fundamental de las diferencias salariales por sexo en España; por ultimo decimos que la existencia de diferencias salariales por sexo sigue siendo un problema real. En España, las mujeres asalariadas perciben un salario equivalente al 68,36 por 100 del de los hombres. El 40 por 100 de la brecha salarial queda explicada por las diferencias en características, mientras que las distintas retribuciones de unas mismas dotaciones son responsables del 60 por 100 restante. En definitiva, una mujer asalariada del sector privado en España cobra, como media, un salario inferior en un 19 por 100 al percibido por un hombre con sus mismas características.

De otro lado el trabajo el autor Rafael Rodriguez Castro en (2014) titulado "DISCRIMINACION SALARIAL POR GÉNERO: CAUSAS Y POSIBLE INFLUENCIA EN LA DEMOGRAFÍA", afirma que la experiencia y el nivel formativo alcanzado de las mujeres en la última década, ha reducido la brecha salarial y son elementos que realmente pueden influir en la existencia de esas diferencias. Donde incluyen la edad afirmando que, en función de ella, la brecha salarial se verá afectada de una manera u otra. Pero a pesar de que la mujer poco a poco va superando en muchos aspectos en lo que se refiere a capital humano respecto al hombre, parece que nunca llega a alcanzar, o por lo menos llega a posicionarse en una situación de igualdad. Sin duda entramos en aspectos sociológicos que afectan, a dos partes del juego. Por un lado, la clase empresarial y por otro la mujer.

Donde da explicación a la clase empresarial, expresando que sigue existiendo ese rechazo, cada vez menor a la incorporación de la mujer en el trabajo y retribuirle con un salario igual por un mismo trabajo. Esto puede ser debido al tipo de cultura que hemos arrastrado durante los años, ya que las personas somos por lo general resistentes a los cambios y seguimos patrones aprendidos a través de la experiencia. De igual forma sigue siendo una barrera para superar el ?coste? que le puede llegar a causar la mujer en lo relativo a que esta se quede embarazada o el ?coste? de que la misma abandone o acorte su vida laboral para el cuidado de su familia y aunque esta tendencia parece que en cierta medida está cambiando, es una realidad que sigue

existiendo hoy en día.

En cuanto a la mujer parece estar ahora mismo entre dos aguas. Por un lado, la mujer quiere incorporarse al mundo laboral, ser independiente o desarrollarse como persona profesionalmente hablando. Y por otro lado parece que quiere seguir un poco el patrón tradicional adoptando el papel de cuidadora de los hijos o la familia. Por lo que la mujer se encuentra en una encrucijada. Aunque se esté adoptando el nuevo rol de mujer independiente en el mundo laboral, parece ser que no termina de desligarse del antiguo y obsoleto papel, sea ya por varios motivos como pueden ser la costumbre o cultura o incluso la presión de una parte importante de la sociedad que sigue existiendo y que prefiere seguir con los métodos tradicionales.

2.1.2. Nacional

A nivel nacional se encuentra el trabajo titulado "DIFERENCIAS SALARIALES POR GÉNERO EN EL DEPARTAMENTO DE SANTANDER, COLOMBIA" realizado por "Alexandra Cortez Aguilar - María Alejandra Flórez vera" en el año 2015 a través del cual se analiza los factores que influyen en el salario por hora diferencias por género para el departamento de Santander durante los años 2012 a 2014. Específicamente se explora si dicha diferencia obedece o no a un factor discriminatorio en el mercado de trabajo santandereano, utilizando los datos proporcionados por la Gran Encuesta Integrada de Hogares GEIH. posteriormente presentaran un análisis descriptivo del mercado laboral santandereano, se realizan estimaciones econométricas aplicando la metodología Blinder-Oaxaca para evidenciar la existencia de un componente discriminatorio. Los resultados indican que, para el total de la muestra en la región santandereana, entre un 25 % y 30 % de las diferencias salariales por género se asocian al componente discriminatorio no explicado; como conclusión de la investigación obtuvieron en términos de la muestra de estudio, de 2012 a 2014 la población masculina ocupada fue de 53,96 % en promedio, y el restante 46,03 % corresponde a población femenina ocupada. Dichas cifras indican que la participación femenina en el mercado laboral se ha incrementado, en comparación a lo que refleja las cifras del 2005 donde la población femenina, como proporción de la población ocupaba, alcanzaba niveles máximos del 20 %; el restante 80 % correspondía a los hombres ocupados, comportamiento que se replica para la mayoría de los rangos de edad (PPMIGS-PS, 2010, p. 23). Durante los tres años, en promedio, las mujeres reciben el 89,28 % del salario que se otorga a los hombres; Los resultados fueron similares a Urdinola y Wodon (2003), puesto que en ocupaciones como empleado doméstico (mujeres) y jornalero (hombres) la masificación de la oferta laboral y el bajo nivel educativo hicieron que la población fuese retribuida con salarios inferiores

2.1.3. Regional

A nivel regional se encuentra el trabajo titulado "DETERMINANTES DE LA BRECHA SALARIAL POR GÉNERO EN COLOMBIA Y EL DEPARTAMENTO DEL HUILA PARA EL AÑO 2014" realizado por CRISTIAN JOSÉ ARIAS BARRERA - LEIDY TATIANA ARIAS IBAÑEZ en el año 2015 en el cual se quiere encontrar cuáles son los determinantes de la brecha salarial por género en Colombia y el departamento del Huila para el año 2014. El propósito de esta investigación es identificar los factores que determinan la brecha salarial que existe entre hombres y mujeres a nivel nacional y departamental. Para ello se utilizó la Misión para el Empleo de las Series de Empleo y Desigualdad (MESEP) suministrada por el DANE, en la cual se recoge información de carácter personal y económico de los hogares que fueron entrevistados.

Una vez organizada la base de datos, se caracterizó cada una de las variables estudiadas, de esta forma fue posible obtener un análisis más detallado de la población objeto de estudio. Es decir, se aplicó la estadística descriptiva para analizar cada una de las variables del estudio, así como la relación entre ellas. Posteriormente se analizó el grado de asociación entre las variables

para proceder a generar el modelo econométrico, la metodología Oaxaca-Blinder, la cual permite tener en cuenta la heterogeneidad existente entre las asignaciones salariales, así como el diferente impacto que tienen los determinantes de los salarios y sus brechas por género en diferentes puntos de la distribución que permitió obtener los resultados finales. algunas de las conclusiones que presenta dicha investigación fueron que la brecha bruta a nivel regional se encuentra por encima de la media nacional con una diferencia de 2 %, influenciado por el hecho de estar ubicado en una zona periférica, al momento de observar la brecha salarial por nivel educativo se evidencia un favorecimiento para el género masculino tanto a nivel nacional como regional las cuales son a nivel de básica primaria es de 12,29 % y 8,29 %, en media es de 33,1 % y 40,7 %, en superior 25,6 % y 32,7 % respectivamente.

2.2. Marco Teorico

La desigualdad salarial entre mujeres y hombres o también diferencia o brecha salarial, alude a la distancia en la retribución media (salario más complementos) de mujeres y hombres.

La discriminación salarial sería la parte de esa diferencia que no queda justificada por una distinta aportación de valor en el desempeño de un trabajo y que solo puede ser explicada en función del sexo de la persona que lo realiza. Este tipo de discriminación es denominada discriminación salarial directa o a posteriori. Así mismo la desigualdad salarial que no es efecto de un factor de discriminación en sí mismo, pero sí consecuencia de las diferencias en el empleo entre mujeres y hombres hace referencia a la teoría denominada la discriminación salarial indirecta o a priori.

Discriminación salarial directa o a posteriori

Es la situación en la que se encuentra una persona que haya sido o pudiera ser tratada en atención a su sexo de manera menos favorable que otra en la misma situación, una de las causas que identifica la discriminación directa es:

- Salarios diferentes
- Despidos por embarazo
- Diferentes contratos

La discriminación directa por razón de sexo, o por cualquier otra condición, está prohibida por la Ley.

Discriminación salarial indirecta o a priori

Dentro de los componentes que determinan la diferencia salarial a priori encontramos:

- **Segregación horizontal:**

mujeres y hombres se concentran en determinados sectores de producción y empleos específicos.

Las mujeres suelen trabajar en sectores en los que su trabajo se valora menos y se paga peor que en los sectores dominados por los hombres. Más del 40 % de las mujeres trabaja

en sanidad, educación y administración pública. Esto es dos veces más que el porcentaje de hombres en los mismos sectores. Si consideramos únicamente los sectores de salud y trabajo social, un 80 % de las personas que trabajan allí son mujeres. Todas las mujeres tuvieron salarios inferiores al de los hombres en todas ellas. El proceso de socialización diferenciado entre mujeres y hombres que se produce en la infancia y juventud también influye en buena manera en nuestra elección de carrera profesional.

■ **Segregación vertical:**

Es cuando se produce esta concentración en determinados grados y niveles de responsabilidad.

La segregación vertical permite abordar el concepto "techo de cristal" una barrera ficticia difícil de traspasar, que impide a la mayoría de las mujeres avanzar en su carrera laboral. La promoción y la permanencia de las mujeres en el mercado laboral está fuertemente condicionada por la existencia de estereotipos, que hace que asuman mayores responsabilidades domésticas y familiares que los hombres.

■ **Trabajo a tiempo parcial:**

Los contratos a tiempo parcial están copados mayoritariamente por mujeres debido a que se ven obligadas a asumir mayores responsabilidades familiares y domésticas.

El trabajo a tiempo parcial, también componente de la diferencia salarial a priori, puede ser una "elección personal", las mujeres son las que más recurren a él para compatibilizar las responsabilidades laborales y familiares, si bien la mayoría preferirían trabajar a tiempo completo.

Inferencia Bayesiana

Antes de comenzar a hablar de inferencia bayesiana debemos hablar acerca del teorema de Bayes. El cual es utilizado para calcular la probabilidad de un suceso, teniendo información de antemano sobre ese suceso.

Podemos calcular la probabilidad de un suceso A, sabiendo además que ese A cumple cierta característica que condiciona su probabilidad. El teorema de Bayes entiende la probabilidad de forma inversa al teorema de la probabilidad total. El teorema de la probabilidad total hace inferencia sobre un suceso B, a partir de los resultados de los sucesos A. Por su parte, Bayes calcula la probabilidad de A condicionado a B.

El teorema de Bayes ha sido muy cuestionado. Lo cual se ha debido, principalmente, a su mala aplicación. Ya que, mientras se cumplan los supuestos de sucesos disjuntos y exhaustivos, el teorema es totalmente válido.

Fórmula del teorema de Bayes

Para calcular la probabilidad tal como la definió Bayes en este tipo de sucesos, necesitamos una fórmula. La fórmula se define matemáticamente como:

$$P[A_n/B] = \frac{P[B/A_n].P[A_n]}{\sum P[B/A_i].P[A_i]}$$

Donde B es el suceso sobre el que tenemos información previa y A(n) son los distintos sucesos condicionados. En la parte del numerador tenemos la probabilidad condicionada, y en la parte de abajo la probabilidad total. En cualquier caso, aunque la fórmula parezca un poco abstracta, es muy sencilla. Para demostrarlo, utilizaremos un ejemplo en el que en lugar de A(1), A(2) y A(3), utilizaremos directamente A, B y C.

Ahora hablaremos de la inferencia bayesiana a continuación:

La inferencia bayesiana está completamente determinada por el uso del teorema de Bayes, la cual nos permite asignar probabilidades a priori, sobre eventos que no son necesariamente de naturaleza aleatoria, es decir que podemos incorporar información extra a nuestro experimento, lo cual nos permite reformular nuestra probabilidad, (probabilidad subjetiva) estas probabilidades se logran afinar mediante el uso del teorema de Bayes; esto no se permite en el contexto de la estadística frecuentista (clásica) donde solo es válido asignar probabilidades, cuando estas se someten a experimentación.

Conceptualmente, la diferencia principal entre la estadística frecuentista y la estadística bayesiana, es la interpretación de la probabilidad.

La Inferencia Bayesiana descansa exclusivamente en el Teorema de Bayes

$$p(\theta | data) \propto p(\theta) p(data | \theta)$$

θ es habitualmente un parámetro, aunque podría ser un valor puntual, una hipótesis o un modelo. Las p son funciones de densidad (o cuantía), $p(\theta)$ es la densidad a priori; $p(data | \theta)$, la verosimilitud (likelihood) de θ aportada por los datos; $p(\theta | data)$ es la densidad a posteriori de θ dados los datos, lo que conduce al Mantra Bayesiano:

El concepto de distribución o densidad a priori es muy controvertido. Algunos estadísticos defienden que en cualquier problema estadístico se puede definir siempre una ley a priori para los datos. Esta distribución representa la información (histórica o subjetiva) del experimentador y se debe trabajar con ella siguiendo las reglas de la probabilidad; podemos hablar también de densidad a priori informativa de la cual podemos tener conocimientos históricos o subjetivos y por otro lado está la densidad a priori no informativa (por ejemplo con densidad o función de masa plana en todos los datos) que se construyen sin usar información a priori y permiten hacer inferencia bayesiana objetiva. Para definir las a veces es necesario recurrir a distribuciones a priori impropias (distribuyen una probabilidad infinita sobre los datos). Pese a su carácter impropio permiten hacer inferencias correctas.

Por otro lado, tenemos La distribución a posteriori nos informa sobre la verosimilitud relativa de que el verdadero valor de μ este en las distintas regiones del espacio paramétrico después de haber observado X_1, \dots, X_n

La densidad a posteriori es proporcional a la densidad a priori por la verosimilitud:

$$p(\theta | data) \propto p(\theta) p(data | \theta)$$

El teorema de Bayes nos dice cómo actualizar nuestras creencias sobre θ a la luz de la evidencia (datos); es un método general de inducción o de 'aprendizaje de la experiencia' a priori→datos→

a posteriori El teorema de Bayes (recordemos que) es un resultado incontrovertible se sigue de los axiomas de Kolmogorov y de la definición de probabilidad condicionada (La estadística bayesiana sí es discutible: la estadística clásica, por ejemplo, no admite que los parámetros sean aleatorios).

3. METODOLOGIA

El objetivo de la metodología es identificar los factores que afectan la desigualdad laboral y de contratación entre hombres y mujeres de la ciudad de Neiva para el año 2016, donde el ajuste del parámetro se realizara por medio de inferencia bayesiana; teniendo en cuenta los datos obtenidos de la Misión para el Empalme de las Series de Empleo, Pobreza y Desigualdad (Mesepe 2016).

3.1. Tipo de Investigación

El tipo de la investigación es de nivel cuantitativo, basado en buscar un modelo con el fin de observar el mejor comportamiento de los datos, y de tipo transversal ya que se centra en la comparación de determinadas características, interés que nace con el fin de implementar una nueva metodología en cuanto a las brechas salariales por genero usando métodos propios de la estadística Bayesiana.

3.2. Población

Para comparar el modelo que mejor se ajusta a los datos de la brecha salarial por género en la ciudad de Neiva, se hizo uso de la base de datos (Misión para el Empalme de las Series de Empleo, Pobreza y Desigualdad) Mesepe en el año 2016, con un total de 778238 datos y 156 variables.

3.3. Diseño Muestral

Para el siguiente trabajo se toma como muestra la población encuestada en la ciudad de Neiva excluyendo las personas que no generan ingresos quedando en total 15193 datos y 4 variables para el año 2016, teniendo en cuenta que 7494 pertenecen al género masculino y 7699 pertenecen al género femenino.

3.4. Instrumentos

Para el desarrollo de la presente investigación se hizo uso de (Misión para el Empalme de las Series de Empleo, Pobreza y Desigualdad) Mesepe en el año 2016-DANE quien es Entidad responsable de la planeación, levantamiento, procesamiento, análisis y difusión de las estadísticas oficiales de Colombia, dichos datos fueron procesados en el Software estadístico R-Studio de donde se obtuvo el análisis descriptivo de la variable de salida, y procedió a la comparación de modelos para escoger el mejor.

3.5. Procesamiento y análisis de información

para este trabajo el procesamiento y análisis de la información se llevaron a cabo las herramientas estadísticas como modelo lineal generalizado.

De acuerdo con Camilo Gómez-Oscar Cerquera-Cristian Orjuela en el año (2017), dicha investigación nos ofrece un acercamiento de la metodología de Oaxaca (1973) y Blinder (1973) descomponiendo la brecha salarial entre hombres y mujeres a través de las funciones de salarios.

$$\ln(w_j) = X_j\beta_j + v_j \quad j = h, m \quad (1)$$

Donde $j = h, m$ indica el grupo el grupo que pertenece a los hombres (h) o las mujeres (m), $\ln(w_j)$ es el logaritmo del salario horario, X_j es la matriz de características observables que intervienen en la productividad, β_j es un vector de coeficientes que presenta los retornos que se

obtienen en el mercado de las características observables en X_j para cada grupo j . Para finalizar v_j es un término no observable que se distribuye con media condicional cero [$E(v_j | X) = 0$] y varianza σ^2 . Entonces al descomponer el logaritmo del salario horario entre hombres y mujeres se obtiene:

$$\ln(w_h) - \ln(w_m) = (X_h\beta_h - X_m\beta_m) + (v_h - v_m) \quad (2)$$

Puesto que el género es una variable binaria que puede tomar el valor de uno si es hombre $D_h = 1$ y cero si es mujer $D_h = 0$, entonces la ecuación de salarios para hombres y mujeres se estima en sus medias muestrales.

$$(\bar{w}_h) = \bar{X}_h\hat{\beta}_h; \quad (\bar{w}_m) = \bar{X}_m\hat{\beta}_m \quad (3)$$

Donde \bar{w}_j es el promedio del salario horario para hombres y mujeres, \bar{X}_j es la media de las características observables y $\hat{\beta}_j$ son los coeficientes estimados. Ahora se construye un término contrafactual, el cual muestra el salario que tendrían las mujeres si fueran remuneradas como los hombres $\bar{X}_m\hat{\beta}_m$, y se obtiene:

$$\bar{\Delta}_w = (\bar{X}_h - \bar{X}_m)\hat{\beta}_h + \bar{X}_m(\hat{\beta}_h - \hat{\beta}_m) \quad (4)$$

El primer componente es la diferencia en el salario promedio de los hombres y el término contrafactual de las mujeres si son remuneradas como los hombres. En otras palabras, es la parte que explica la brecha salarial por las diferencias en las características productivas entre hombres y mujeres y que es conocido como efecto característico. Por su parte, el segundo componente es la diferencia en el ingreso promedio contra fáctico (ingreso promedio que recibirán los hombres si tuvieran las características productivas de las mujeres) y el salario promedio observado de las mujeres, ya que solo interesa la diferencia de la remuneración en las características observables (Debido a que se controlan las diferencias en las características productivas), este componente se conoce como efecto remuneración y evalúa la posible discriminación salarial.

No obstante, debido a su rol de madre, las mujeres acceden al mercado siempre y cuando su salario sea mayor al salario del mercado, el cual le permita cubrir el costo de oportunidad en el desempeño de labores de madre y ama de casa, lo cual genera un sesgo de selección. Para corregir este sesgo se aplica la metodología de Heckman (1979). Primero se estima la probabilidad de que el salario de los individuos se explique al utilizar como instrumento la tenencia de hijos en el hogar menores de 18 años junto al resto de características observables.

$$Pr(w_{h,m} | X_{h,m}) = \theta(X'_{h,m}\delta) \quad (5)$$

Una vez calculada en la primera etapa la probabilidad del salario, junto a las características que influyen en el rol de madre, se estima la inversa de la razón de Mills (Cociente entre la función de densidad y la función de densidad acumulada de la normal estándar) $\lambda(h, m)$ la cual se introduce en la segunda etapa como una variable explicativa adicional que corrige el sesgo de selección.

$$\hat{\Delta}_w - (\hat{\gamma}_h\lambda_h - \hat{\gamma}_m\lambda_m) = (\bar{X}_h - \bar{X}_m)\hat{\beta}_h + \bar{X}_m(\hat{\beta}_h - \hat{\beta}_m) \quad (6)$$

No obstante, si se quiere analizar la descomposición del diferencial salarial en la distribución incondicional (como los cuantiles) del ingreso laboral, y si se desea saber qué puede pasar con los efectos característica y remuneración, es decir, poder aislar ambos efectos de la distribución no condicional, se utilizan en principio la regresión por cuantiles y las metodologías de descomposición desarrolladas por Machado y Mata (2004), Autor, Katz y Kearney (2005), Melly (2005) y Firpo, Fortiín y Lemieux (2011). En este caso, el método de Melly (2005) surge como alternativa para estimar los efectos característica y remuneración en la distribución incondicional del

salario, al tener en cuenta el problema de sesgo por selección en el ingreso laboral de las mujeres. Para ello, propone un estimador no condicional que puede generarse al integrar la distribución condicional en todo el rango de la distribución de regresores.

Sin embargo, cuenta con un problema de falta de mono tonicidad, es decir, $\theta_j \leq \theta_k$ no necesariamente implica $X_i\hat{\beta}(\theta_j) \leq X_i\hat{\beta}(\theta_k)$. Para evitar esto, propone un cambio de variable en el proceso de integración. Dado esto y a través de Koenker y Bassett (1978) se obtiene:

$$F_{(Y|X)}^{-1}(\theta | X) = X_j\beta(\theta) \quad (7)$$

Donde $F_{(Y|X)}^{-1}(\theta | X)$ es el θ -esimo cuantil de Y condicional sobre X. Como se sabe, la relación lineal entre los cuantiles de Y y X es similar a la relación lineal entre las variables explicativas y explicada bajo Mínimos Cuadrado Ordinarios. Entonces al reemplazar $F_{(Y|X)}^{-1}(\theta | X)$ por su estimador consistente $X_j\beta(\theta)$ y al tomar el ínfimo del conjunto de soluciones finitas sin que sea la única solución se obtiene:

$$\hat{a}(\hat{\beta}, X) = \mathit{infa} : \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^J (\theta_j - \theta_{j-1}) 1(X_j\beta(\theta) \leq a) \geq \tau \quad (8)$$

Donde a es consistente y distribuido como una normal, por su parte la dificultad de estimar la varianza asintótica conlleva a utilizar el método de Bootstrap para hacer inferencia estadística. La descomposición de los diferenciales en la distribución del salario para hombres y mujeres se realiza a partir de la distribución incondicional. Los estimadores obtenidos se utilizan para poder generar el ingreso contra fáctico en ambos géneros, ya que la idea principal de este estimador es la posibilidad de simular las distribuciones hipotéticas que se puedan utilizar para descomponer las diferencias en la distribución.

Para obtener el efecto en los cambios de las características, remuneración (incluye el efecto residual), se debe primero calcular la distribución contrafactual del salario de las mujeres y suponer que la distribución de sus características observables ha sido como la de los hombres. Entonces, al minimizar (8) para la distribución de los hombres y a través de los coeficientes estimados para las mujeres, se genera:

$$\hat{a}(\beta_m, X_m) - \hat{a}(\beta_h, X_h) = [\hat{a}(\beta_{\text{mediam},rh,X_m}) - \hat{a}(\beta_h, X_m)] + [\hat{a}(\beta_h, X_m) - \hat{a}(\beta_h, X_h)] \quad (9)$$

Para corregir por sesgo de selección en la distribución del salario horario se sigue a Buchinsky (1998), dado esto se utiliza la estimación paramétrica de la inversa de la ratio de Mills y su valor cuadrado, los cuales se agregan como variables adicionales al método de descomposición del diferencial salarial por género propuesto por Melly (2005). Esto permitirá corregir el sesgo de selección debido al rol de madres de las mujeres en cada cuantil de la distribución salarial.

$$\hat{a}(\beta'_m, X'_m) - \hat{a}(\beta'_h, X'_h) = [\hat{a}(\beta'_{\text{mediam},rh,X'_m}) - \hat{a}(\beta'_h, X'_m)] + [\hat{a}(\beta'_h, X'_m) - \hat{a}(\beta'_h, X'_h)] \quad (10)$$

MODELO LINEAL GENERALIZADO

De acuerdo a (María Durbán) en el trabajo titulado 'Modelos Lineales Generalizados' Gran cantidad de datos son recogidos por científicos, médicos, o empresas como datos de conteo. En general, este tipo de datos aparecen en cuatro formatos distintos:

- Datos como frecuencias, donde contamos cuantas veces ocurre algo, pero no sabemos cuántas veces no ocurre.

- Datos como proporciones, donde ambos, el número de veces que ocurre algo, y el tamaño total del grupo es conocido.
- Datos por categorías, donde la variable cuenta cuantos individuos hay en cada nivel de una variable categórica.
- Datos binarios, donde recogemos la presencia o ausencia de una característica.

Hay cuatro razones por las que es erróneo utilizar un modelo de regresión normal para datos de conteo:

1. Puede dar lugar a predicciones negativas.
2. La varianza de la variable respuesta no es independiente de la media.
3. Los errores no siguen una distribución Normal.
4. Los ceros que aparecen en la variable respuesta dan problemas a la hora de transformar las variables.

Sin embargo, si los datos son elevados, es posible utilizar la distribución Normal.

Familia Exponencial

Un concepto importante que unifica todos los GLMs es la familia exponencial de distribuciones. Este concepto fue presentado por primera vez en Fisher (1934).

Todas las distribuciones pertenecientes a la familia exponencial tienen una función de densidad (o de probabilidad) que se puede expresar de la siguiente forma:

$$f(y; \theta, \phi) = \exp \left\{ \frac{y\theta - b(\theta)}{a\phi} + c(y, \phi) \right\} \quad (11)$$

donde, en cada caso, $a(\cdot)$, $b(\cdot)$ y $c(\cdot)$ serán funciones específicas. El parámetro θ es lo que se llama parámetro canónico de localización y ϕ es un parámetro de dispersión. La distribución Binomial, Poisson y Normal (entre otras) son miembros de la familia exponencial. El caso más importante es el de la distribución Normal, cuya función de densidad es:

$$f(y; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} \exp \left\{ -[y - \mu]^2 / 2\sigma^2 \right\}$$

que podemos reescribir como

$$= \exp \left\{ (y\mu - \mu^2/2) / \sigma^2 - \frac{1}{2} [y^2 / \sigma^2 + \ln(2\pi\sigma^2)] \right\}$$

Por lo tanto,

$$\theta = \mu, b(\theta) = \mu^2/2, a(\phi) = \phi, \phi = \sigma^2$$

y

$$c(y, \phi) = \frac{1}{2} \left[\frac{y^2}{\sigma^2} + \ln(2\pi\sigma^2) \right]$$

Componentes de un modelo lineal generalizado

Comencemos con el modelo de regresión estándar:

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon \quad \varepsilon \sim N(0, \sigma^2 I) \quad E(y) = \mu = \mathbf{X}\beta$$

donde $X\beta$ es una combinación lineal de las variables predictoras llamada predictor lineal (el cual se representa como η), en este caso la media μ está directamente relacionada con el predictor lineal, ya que en este caso $\mu = \eta$. Usando este modelo sencillo, podemos ver que hay dos componentes en el modelo: la función de probabilidad de la variable respuesta y la estructura lineal del modelo. En general, un modelo lineal generalizado tendrá los siguientes componentes:

1. **Componente aleatorio:** y es un vector aleatorio procedente de una distribución que pertenece a la familia exponencial y cuya media es μ .
2. **Componente sistemático:** es el predictor lineal $\eta = X\beta$.
3. **La función link:** es una función monótona, derivable que establece la relación entre la media y el predictor lineal.

$$\eta = g(\mu) \quad E(y) = \mu = g^{-1}(\eta)$$

En el caso del modelo de regresión ordinaria, $\mu = \eta$, por lo tanto la función link es la identidad.

Hay muchas opciones para la función link. La función link canónica es una función que transforma la media en el parámetro canónico θ

$$\eta = g(\mu) = \theta \Rightarrow g$$

Estimación del parámetro de dispersión

Con la excepción de la variable Binomial y Poisson, el parámetro de dispersión no tiene por qué ser conocido, y tendrá que ser estimado. Cuando estimamos β , no es necesario conocer ϕ , ya que es sistema de ecuaciones de la función score, es independiente de ϕ . En el caso de la distribución Normal, el parámetro de dispersión $\phi = \sigma^2$ es estimado de modo que la varianza residual escalada es igual a los grados de libertad, es decir, $\hat{\phi}^2 = RSS/(n-p) = D/(n-p)$, y por lo tanto $d/\hat{\phi} = n-p$. La extensión de esto al caso de los GLMs sería estimar ϕ como la media de los residuos de Pearson al cuadrado,

$$\hat{\phi} = \frac{\sum_{i=1}^n r_i^2 P}{n-p}$$

Inferencia

Hay varias maneras de hacer contrastes de hipótesis sobre las componentes de β .

1. Si queremos contrastar por ejemplo, $\beta_1 = 0$ (el primer componente de β), el procedimiento sería calcular $\hat{\beta}_1$ y $se\hat{\beta}_1$ y comparar $|\hat{\beta}_1|/se\hat{\beta}_1$ con $N(0, 1)$, y rechazar $\beta_1 = 0$ si esa cantidad es demasiado grande.
recuerda que $se(\hat{\beta}_1)$ es la raíz cuadrada del elemento (1, 1) de la inversa de la matriz

$$\frac{\partial^2 l}{\partial \beta \partial \beta'} \Big|_{\beta}$$

De modo que aquí estamos usando la normalidad asintótica del estimador máximo verosímil $\hat{\beta}$, y la fórmula de su varianza asintótica.

2. Si la muestra de datos es pequeña, la aproximación a la Normal, puede no funcionar bien. Una alternativa es utilizar el Análisis de la Varianza. La base de este método es utilizar una medida del ajuste del modelo que mida la discrepancia entre los datos ajustados por el modelo y los datos. Para datos normales esta medida es la Suma de cuadrados residual, para datos no-Normales es el Deviance. En general, el deviance se basa en valor (maximizado) de la verosimilitud. El test se basa en la reducción de esta medida de ajuste del modelo cuando incluimos una nueva variable. Una buena aproximación para la distribución que sigue esta reducción de la verosimilitud es la distribución Chi-cuadrado. Por lo tanto, otra forma de hacer contrastes sobre los parámetros es mediante un test Chi-cuadrado.

En general, supongamos que queremos contrastar $\beta \subset \omega_r$ (el modelo reducido, es decir, $\beta_1 = 0$) frente a $\beta \subset \omega_f$ (el modelo completo) donde $\omega_r \subset \omega_f$. Consideremos la diferencia entre los máximos del logaritmo de la función de verosimilitud para el modelo completo y el reducido. Sea L_f el valor maximizado del logaritmo de la verosimilitud bajo el modelo completo (todos los β_1 presentes) y sea L_r el valor maximizado del logaritmo de la verosimilitud bajo el modelo reducido (con $\beta = 0$). Es claro que L_f es al menos tan grande como L_r . El test estadístico es:

$$S(\omega_r, \omega_f) = -2(L_r - L_f) = 2 \sum_{i=1} [y_i(\tilde{\theta}_i - \hat{\theta}_i) - (b(\tilde{\theta}_i) - b(\hat{\theta}_i))]/\phi$$

donde $\tilde{\theta}_i = eml$ (estimador máximo verosímil) bajo ω_r y $\hat{\theta}_i = eml$ bajo ω_f .

La distribución de $S(\omega_r, \omega_f)$ se aproxima por una X^2 con grados de libertad igual a la diferencia entre el número de parámetros en el modelo completo y el reducido (si estamos contrastando $\beta_1 = 0$, entonces, X_1^2). Los intervalos de con-

anza para los parámetros se calculan también usando la distribución X^2 .

Deviance escalado: Un modelo saturado, es un modelo que ajusta perfectamente los datos (es decir, que tiene tantos parámetros como observaciones y los valores ajustados son iguales a los observados). Sea l_s el logaritmo de la verosimilitud de modelo saturado, y sea l_m el valor maximizado del logaritmo de la verosimilitud del modelo de interés. El Deviance escalado del modelo de interés es

$$Deviance\ escalado = S(y, \hat{\mu}) = -2(L(\hat{\mu}, \phi, y) - L(y, \phi, y)) = -2(l_m - l_s)$$

y el deviance es

$$Deviance = D(y, \hat{\mu}) = S(y, \hat{\mu})\phi$$

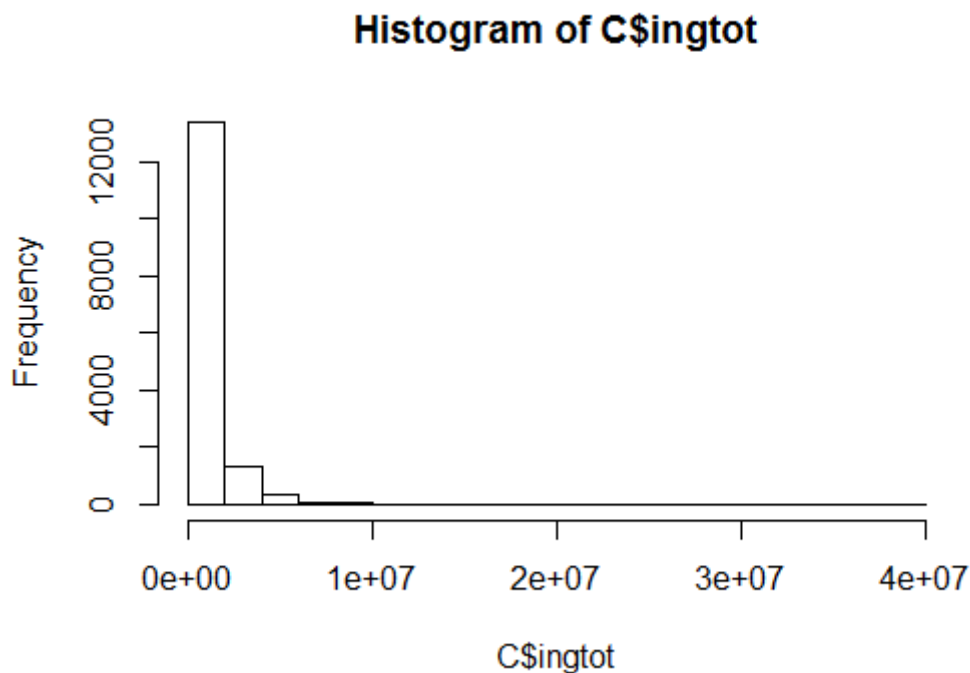
Por lo tanto, el deviance escalado se puede ver como un test estadístico para contrastar la hipótesis de que los parámetros del modelo saturado (y que no está en el modelo de interés) son iguales a 0. En el caso de datos de Poisson y Binomial, $\phi = 1$, y por lo tanto ambos deviance coinciden. En el caso de la Normal, $\phi = \sigma^2$, entonces, el deviance escalado sigue una distribución F .

4. ANÁLISIS DE RESULTADOS

Análisis descriptivo respecto a los ingresos totales

A través de los datos trabajados encontramos que los datos no presentan una distribución normal, pues se observa un nivel muy alto en las personas que reciben un ingreso salarial bajo.

Figura 1: Diagrama de barras de los ingresos salariales por género.

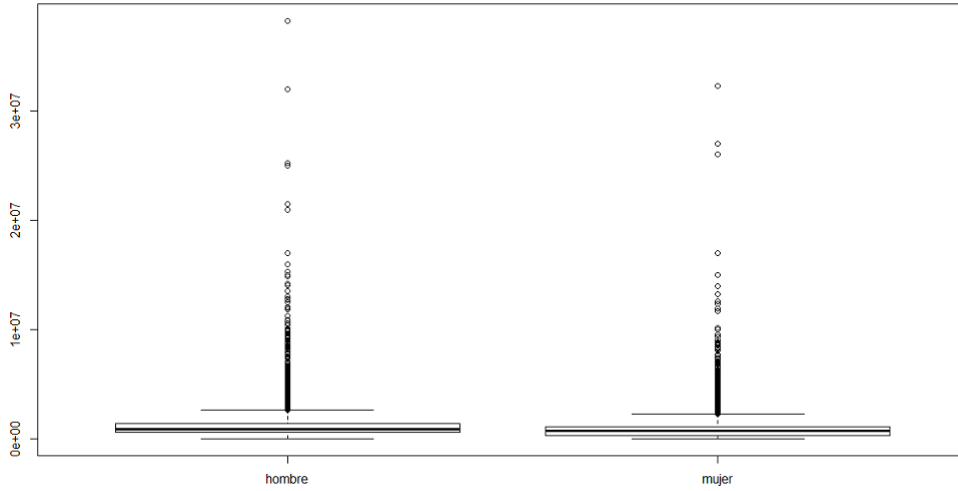


El grafico anterior observamos que la distribución de los datos presenta una asimetría positiva, presentando un comportamiento que describe el 88,3% de las personas se encuentran ganando un salario entre 0 y \$ 2,000,000, y el 11,7% de las personas generan ingresos superiores a \$ 2,000,000 llegando a ganar \$ 38,231,133.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
3333	400000	785454	1090296	1210000	38231133

A partir de la información obtenida de 15,193 personas encuestadas en la ciudad de Neiva, se puede concluir que el promedio de ingresos anuales declarados es de \$ 1,090,296; igualmente se observa que existe una gran dispersión en la distribución de los ingresos \$ 3,333 como mínimo y \$ 38,231,133 como máximo. Se puede afirmar, a partir de los datos que el 25 por ciento de los encuestados tienen como ingresos máximos de \$ 400,000 y el 50 por ciento tiene ingresos iguales o superiores a \$ 785,454, de igual forma se manifiesta que coeficiente de asimetría es de $7,143 > 0$ indicando un bajo ingreso salarial, de otro lado encontramos que el coeficiente de curtosis es $105,487 > 0$ lo cual me explica que los datos son homogéneos.

Figura 2: Diagrama de Cajas de Genero con respecto al salario.



La gráfica de caja muestra los ingresos salariales por género. Indicando una similaridad de los ingresos salariales medios por género, pero dentro los ingresos salariales de los hombres se percibe una alta variabilidad indicando ingresos salariales bajos, con respecto al de las mujeres que presenta una baja variabilidad, al igual se evidencia que los hombres generan mejores ingresos que las mujeres pues la presencia del dato atípico nos refleja que al menos un hombre tiene ingresos por encima de \$ 38,000,000.

Modelo Matemático

La importancia de los modelos matemáticos es con el fin de hacer uso de las fórmulas matemáticas para representar la relación entre distintas variables, parámetros y restricciones, por tanto, un modelo matemático es una representación simplificada, a través de ecuaciones, funciones o fórmulas matemáticas, de un fenómeno o de la relación entre dos o más variables, así el modelo a utilizar será:

Modelo 1: Distribución Normal

$$y = \beta_1 + \beta_2x_2 + \dots \beta_ix_i + \varepsilon \tag{12}$$

con un comportamiento Normal $y \sim N(\mu, \sigma^2)$

Modelo 2: Distribución Gamma

$$y = \beta_1 + \beta_2x_2 + \dots \beta_ix_i + \varepsilon \tag{13}$$

con un comportamiento Gamma $y \sim Gamma(\alpha, \beta)$

Dado que en la literatura se han realizado estudios de la brecha salarial por género, todos estos estudios se realizan desde el punto de vista de la estadística clásica, en este sentido se propone un modelo lineal generalizado, tomando como variable de salida los ingresos. desde el punto de vista bayesiano se propone que la variable dependiente sigue una distribución de la familia gamma, es decir, $y \sim Gamma(\alpha, \beta)$, entre los GLM encontramos el log el cual se especifica en el siguiente modelo con el fin de dar respuesta al estudio que se lleva a cabo:

	β_0	β_1	β_2	Residual desviansa	AIC	Kurtosis
Modelo 1	1081545	26851	-298414	$3,0270e + 16$	473395	110,1293
Error	24796	3043	22933			
Modelo 2	13,844404	0,029939	-0,268101	15585	452417	5,616975
Error	0,022355	0,002744	0,020676			

Cuadro 1: Modelo 1 Vs Modelo 2

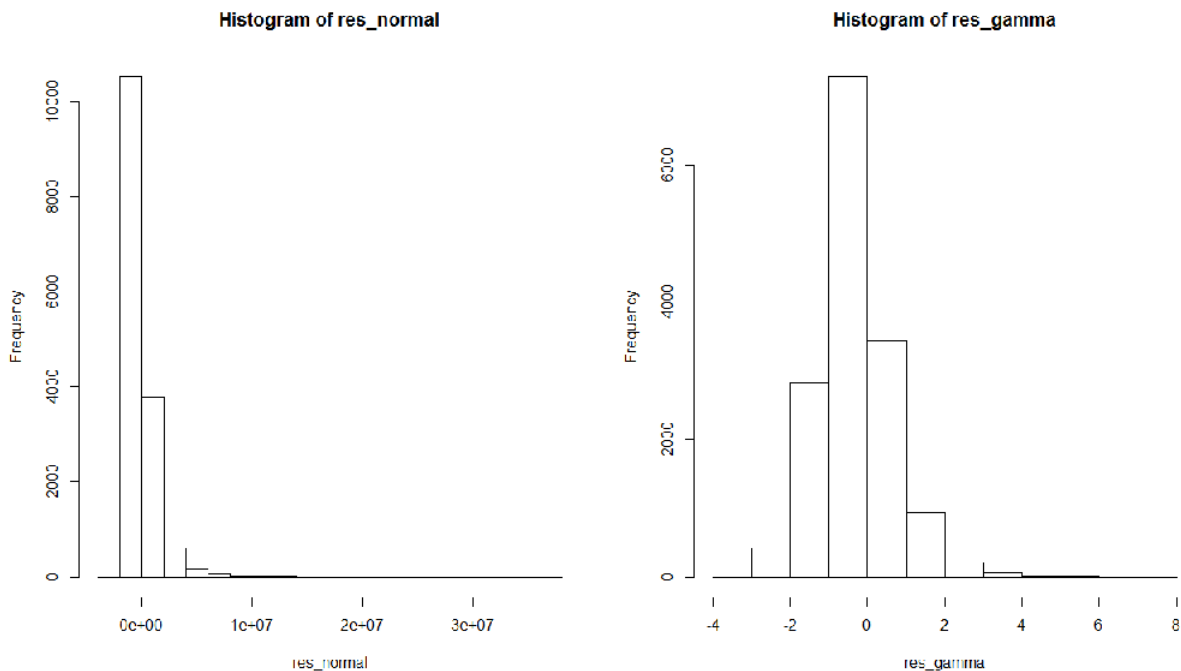
Interpretación del modelo

El parametro β_0 en el modelo de distribución normal nos explica que la probabilidad que un hombre genere mejores ingresos es de 1081545 más que la mujer, mientras que en el modelo de distribución Gamma la probabilidad es de 13,94 unidades mas que la mujer.

El modelo de distribución normal explica que por cada nivel de estudio se incrementa 26851 unidades con respecto a los ingresos salariales y que el ingreso de las mujeres es -298414 unidades menos que el de los hombres, a diferencia de la distribución Gamma que me explica que por cada nivel de estudio se incrementa en 0,029 unidades con respecto a los ingresos salariales y las mujeres pierden -0,26 unidades a diferencia de los hombres.

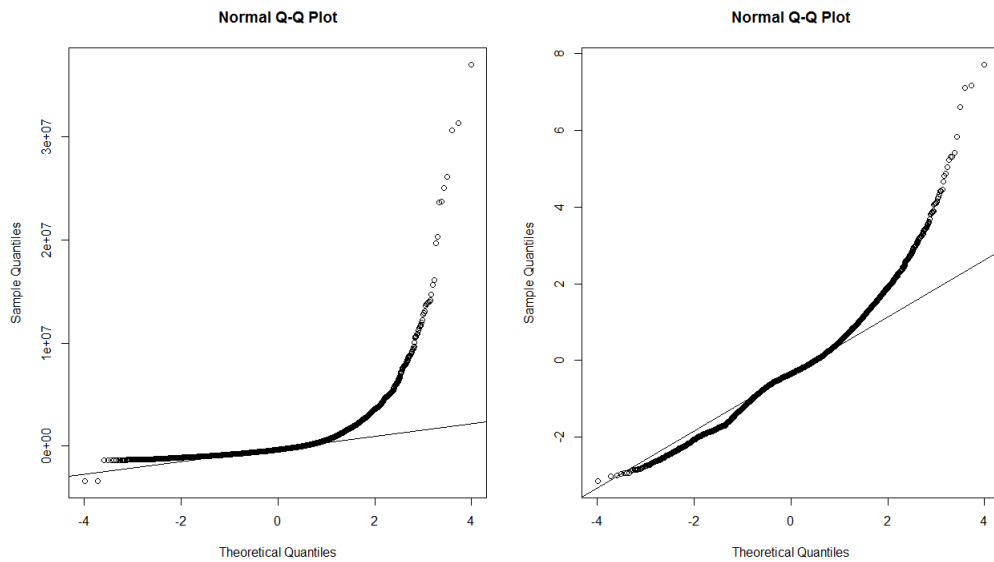
Dentro de los valores de curtosis arrojados en los dos modelos se evidencia que para la distribución normal el valor es de 110,1293 presentando un comportamiento no normal, y ajustado a una distribución Gamma que me permite trabajar con los datos originales y hacer un ajuste a los parámetros de salida, se observa un valor de 5,6169 lo que explica que los datos tienen un comportamiento normal, de igual forma la presencia del valor de AIC no es una prueba del modelo en el sentido de la prueba de hipótesis, sino que es una prueba entre los modelos, es una herramienta para la selección del mejor modelo indicando el valor mínimo entre modelos, por tanto en la distribución normal el valor del AIC es 473395 y en la distribución Gamma es 452417, presentando una diferencia significativa entre ambos modelos, llegando a elegir la Distribución Gamma como el modelo que mejor se ajusta.

Figura 3: Diagrama de Barras de los Residuales Normales Vs Residuales Gamma



De la gráfica anterior se evidencia que los residuales de la distribución normal presentan un comportamiento asimétrico positivo indicando valores muy bajo (a la izquierda) y valores muy altos (a la derecha), de igual forma presenta una distribución con mayor curtosis, mientras que en el diagrama de barras de los residuales de la distribución gamma se evidencia una normalización de los datos; concluyendo que la distribución Gamma es el modelo que mejor se ajusta para evidenciar el comportamiento de los datos de una manera normal.

Figura 4: Curva de normalidad de Distribución Normal Vs Distribución Gamma.



Del Grafico anterior se observa una comparación de los cuantiles (sample Quantiles) con los los cuantiles teóricos (Theoretical Quantiles) de una distribución normal Vs distribución Gamma, donde los datos generados no siguen una distribución normal estándar (rnorm) La linealidad de los puntos sugiere que los datos se distribuyen normalmente evidenciándose como los puntos se alejan de la recta, especialmente a la derecha del gráfico, presentando una asimetría positiva, la cola a la derecha de la media es mas larga que a la izquierda, de otro lado encontramos que los datos generados por la distribución Gamma se ajustan mejor a la recta indicando que es el mejor ajuste al modelo.

5. CONCLUSIONES

- En el trabajo se logro evidenciar que al ajustar un modelo bayesiano la diferencia salarial por genero presenta una variación grande con respecto al ajuste de un modelo gaussiano, dado que AIC de la distribución normal es 473395 y en la distribución Gamma es 452417, presentando una diferencia significativa entre ambos modelos.
- De acuerdo con la exploración se observó que las transformaciones de los datos originales no siempre consiguen corregir la falta de normalidad, la heterocedasticidad (varianza no constante) o la no linealidad de nuestros datos. Por ello una de las opciones a la transformación de la variable dependiente/respuesta y a la falta de normalidad es el uso de los modelos lineales generalizados.
- Los modelos lineales generalizados son una extensión de los modelos lineales que permiten utilizar distribuciones no normales de los errores (binomial, Poisson, gamma, binomial no negativa, etc) y varianzas no constantes, entonces, cuando la variable respuesta/dependiente se aproxime a una distribución Gamma, la varianza se incrementa de una manera no lineal con respecto a la media.
- Modificar una función de densidad o de probabilidad hacia la familia exponencial permite de una manera sencilla evaluar los estadísticos suficientes, y con ello determinar los estadísticos que permiten estimar los valores de los parámetros.
- Finalmente, nuestros resultados nos permiten concluir que es posible obtener estimaciones de muy buen rendimiento para el modelo cuando a la variables respuesta se asume una distribución previa como a priori.

6. RECOMENDACIONES

se recomienda el uso de la distribución Gamma para el análisis de datos de esa naturaleza debido a que no se requiere de transformar los datos originales si no de transformar sus residuales; adicionalmente tambien se recomienda un estudio que contemple las deficiencias del mercado laboral tales como las derivadas de la sub-calificación y sobre-calificación en distintas ocupaciones y por ultimo se recomienda a las empresas del sector tanto publico como privado la realización de capacitaciones constantes al personal sobre el derecho a la igualdad.

7. BIBLIOGRAFIA

Referencias

- [1] ARANGO, Luis Eduardo y POSADA, Carlos Esteban. La participación laboral en Colombia. Colombia. Banco de la República. 2002.
- [2] GOMEZ, Guadalupe y DELICADO, Pedro. Curso de inferencia y decisión. Cataluña - España. Universidad Politecnica de Cataluña. 2016.
- [3] PRECIADO, Ginna Paola. Determinantes de la oferta laboral femenina en Colombia: caso comparado entre el tercer trimestre del 2009 y 2013. Santiago de Cali-Valle del Cauca. Universidad del Valle. 2015.
- [4] TABÓN, Herrera Camilo y RODRÍGUEZ, Vargas Frey León. Factores que determinan la probabilidad de participación laboral en el área metropolitana de Medellín. Medellín - Colombia. Universidad EAFIT. 2015.
- [5] MARCILLO, Yépez Edgar y ZAMBRANO, Juan Carlos. Determinantes de la participación en el mercado de trabajo. Un estudio para el área metropolitana de Pasto. Pasto-Nariño. Universidad de Nariño. 2010.
- [6] MESA, Páez Lesley Ofelia, RIVERA, Lozano Miller y ROMERO, Davila Jesús Andrés. Descripción general de la inferencia Bayesiana y sus aplicaciones en los procesos de gestión. Bogotá - Colombia. Universidad del Rosario. 2011.
- [7] ALÁEZ, Aller Ricardo y ULLIBARRI, Arce Miren. Discriminación salarial por sexo: Un análisis del sector privado y sus diferencias regionales en España. España. Universidad Pública de Navarra. 2001.
- [8] GÓMEZ, Segura Camilo Fabian; CERQUERA, Losada Óscar Hernán y ORJUELA, Yacue Cristian Felipe. Evolución de la brecha salarial por género en Colombia 2004-2012; un análisis del fenómeno techo de cristal. Medellín - Antioquia. Universidad de Medellín. 2018.
- [9] BRAVO, Barriga Diana Elizabeth. Brecha remuneracional de género en el segmento profesional de la gran minería chilena: Un estudio de caso. Santiago de Chile - Chile. Universidad de Chile. 2017.
- [10] DURBÁN, María. Modelos Lineales Generalizados.
- [11] ARIAS, Barrera Cristian José y ARIAS, Ibañez Leidy Tatiana. Determinantes de la brecha salarial por género en Colombia y el departamento del Huila para el año 2014. Neiva - Huila. Universidad Surcolombiana. 2015.
- [12] GALVIS, Luis Armando. Diferencias salariales por género y región en Colombia: Una aproximación con regresión por cuantiles. Cartagena de Indias - Bolívar. Banco de la República. 2010.
- [13] BARBOSA, Wilson Julián y CASTAÑEDA, Beltrán Johan. Determinantes de las diferencias salariales y laborales por género, un análisis econométrico para Colombia entre los años 2005 - 2014. Bogotá - Colombia. Universidad de la Salle. 2017.