



САНХҮҮ ЭДИЙН ЗАСГИЙН ИХ СУРГУУЛЬ
ӨДРИЙН ХӨТӨЛБӨР
ЭКОНОМИКСИЙН ТЭНХИМ

Булгадар
Жамсрангийн БАТХИШИГ

ДИСКРИМИНАНТ ШИНЖИЛГЭЭНИЙ АРГА ЗҮЙ,
ТҮҮНИЙ ХЭРЭГЛЭЭ



Мэргэжлийн индекс
031101

Эдийн засгийн ухааны бакалаврын
зэрэг горилсон
ДИПЛОМЫН ТӨСӨЛ

Удирдсан
П. Гантөмөр /Ph.D/

Улаанбаатар. 2019



САНХҮҮ ЭДИЙН ЗАСГИЙН ИХ СУРГУУЛЬ
ЭКОНОМИКСИЙН ТЭНХИМ



Булгадар
Жамсрангийн БАТХИШИГ

ДИСКРИМИНАНТ ШИНЖИЛГЭЭНИЙ АРГА ЗҮЙ,
ТҮҮНИЙ ХЭРЭГЛЭЭ



Мэргэжлийн индекс
031101

Эдийн засгийн ухааны бакалаврын зэрэг
горилсон дипломын төсөл

Удирдагч: П. Гантөмөр /Ph.D/

Шүүмжлэгч: Д. Гансүлд /MA/

Улаанбаатар. 2019

УДИРТГАЛ

Сэдвийн нэр: *Дискриминант шинжилгээний арга зүй, түүний хэрэглээ*

Товч танилцуулга: Энэхүү дипломын ажлын хүрээнд дискриминант шинжилгээний арга зүй, хэрэглээг дэлгэрэнгүй авч үзлээ. Энэ шинжилгээний арга зүй нь чанарын үл хамаарах хувьсагчдыг ашиглан хамааран хувьсагчдад таамаглал хийн ангилдгаараа давуу талтай арга зүй билээ. Дискриминант шинжилгээ дотроо шугаман дискриминант шинжилгээ, хэвийн дискриминант шинжилгээ, дискриминант шинжилгээний шинжилгээ хэмээх бүрдэл хэсгүүдтэй байдгаас шугаман дискриминант шинжилгээ судалгааны ажлуудад түлхүү хэрэглэж байна. Тус шинжилгээ нь вариацийн шинжилгээ (ANOVA), логистик регрессстэй төстэй мэт боловч функц болон хэмжүүрүүдийн шугаман хослолоосоо хамаараад ялгаатай үр дүн гарах нь бий. Судалгааны ажлын хүрээнд онол арга зүйн олж авсан мэдлэг дээр тулгуурлан оюутны академик гүйцэтгэлийг таамаглах шинжилгээ хийв.

Эдийн засгийн бүтээлийн сэтгүүлийн ангиллын индекс: *C11*

Түлхүүр үгс: *Дискриминант шинжилгээ, дискриминант функц, шугаман дискриминант шинжилгээ, оюутнуудын гүйцэтгэл*

АГУУЛГА

УДИРТГАЛ	i
ОРШИЛ	1
I БҮЛЭГ. ДИСКРИМИНАНТ ШИНЖИЛГЭЭНИЙ АРГА ЗҮЙН ТАЛААРХ СУДЛАГДСАН БАЙДЛЫН ТОЙМ	3
1.1 Оюутнуудын амжилтыг дискриминант шинжилгээ ашиглан хийсэн судалгааны ажлууд.....	3
1.2 Зах зээл, хөгжлийн асуудлуудын мэдээллийг ашиглан олон улсын дискриминант шинжилгээ хийсэн судалгааны ажлууд	5
1.3 Банкны мэдээллийг ашиглан дискриминант шинжилгээ хийсэн ажлууд	7
1.4 Хэрэглэгчийн худалдан авалттай холбоотой дискриминант шинжилгээг ашиглан хийсэн судалгааны ажлууд.....	8
1.5 Дискриминант шинжилгээний хөгжүүлэлттэй холбоотойгоор хийгдсэн судалгааны ажлууд	8
II БҮЛЭГ. ДИСКРИМИНАНТ ШИНЖИЛГЭЭНИЙ ОНОЛ, АРГА ЗҮЙ	10
2.1 Бүлгийг тусгаарлах нь	10
2.1.1 Хоёр бүлгийн дискриминант функц	11
2.1.2 Хоёр бүлгийн дискриминант шинжилгээ ба олон хүчин зүйлсийн регрессийн хамаарал	12
2.2 Олон бүлгийн дискриминант шинжилгээ.....	12
2.2.1 Дискриминант функцүүд	12
2.2.2 Дискриминант функцийн хамаарлын хэмжүүр	14
2.2.3 Стандартчилагдсан дискриминант функцүүд.....	15
2.3 Ач холбогдлын тест	16
2.3.1 Хоёр бүлгийн тохиолдлын тест.....	16
2.3.2 Олон бүлгийн тохиолдлын тест	16
2.4 Дискриминант функцийн тайлбар.....	18
2.4.1 Стандартчилагдсан коэффициент	19
2.4.2 Хэсгийн F -утгууд.....	20
2.5 Хувьсагчид болон дискриминант функцийн хоорондох хамаарал.....	21

2.6 Цэгэн Түгэлт	21
2.7 Хувьсагчдыг сонгох шатлал	22
III БҮЛЭГ. ЭМПИРИК СУДАЛГАА	23
3.1 Дискриминант шинжилгээ болон бусад шинжилгээнүүдээр хийсэн ангиллын үр дүн	26
ДҮГНЭЛТ, САНАЛ	28
НОМ ЗҮЙ	30
ХАВСРАЛТ	32

ХҮСНЭГТҮҮДИЙН ЖАГСААЛТ

Хүснэгт 1: Хувьсагчдын мэдээлэл	23
Хүснэгт 2: Тодорхойлогч статистик үзүүлэлтүүд	24
Хүснэгт 3: Түүврийн дундаж векторууд	24
Хүснэгт 4: Боксын М Тест	25
Хүснэгт 5: Ач холбогдлын тест	25
Хүснэгт 6: Хотеллинг Т квадрат тест	26
Хүснэгт 7: Дискриминант шинжилгээний үр дүн	26
Хүснэгт 8: Дискриминант шинжилгээний ангиллын хувь	26
Хүснэгт 9: Логистик регрессийн ангиллын үр дүн	27
Хүснэгт 10: Логистик регрессийн ангиллын хувь	27

ХАВСРАЛТУУДЫН ЖАГСААЛТ

Хавсралт 1: Шинжилгээний хэсэгт ашигласан өгөгдөл	32
Хавсралт 2: Өгөгдөлд тулгуурлан хийсэн хүчин зүйлсийн шинжилгээ	33
Хавсралт 3: Өгөгдөлд тулгуурлан хийсэн коварианс болон корреляцийн матрицууд ...	35
Хавсралт 4: Өгөгдөлд тулгуурлан хийсэн нэг чиглэлт ANOVA-н үр дүн	36
Хавсралт 5: Өгөгдөлд тулгуурлан хийсэн дискриминант шинжилгээний ангиллын үр дүн	39
Хавсралт 6: Өгөгдөлд тулгуурлан хийсэн логистик регрессийн үр дүн	39
Хавсралт 7: Өгөгдөлд тулгуурлан ангиллын шинжилгээгээр тооцоолсон үр дүн	40

ТОВЧИЛСОН ҮГС, НЭР ТОМЪЁОНЫ ТАЙЛБАР

- ДШ-Дискриминант шинжилгээ
- ИХБТХӨ-Их хэмжээст, бага түүврийн хэмжээтэй өгөгдөл
- ӨБШ-Өгөгдлийн бүрэлдэхүүний шинжилгээ
- ҮБШ-Үйлдвэрлэлийн боломжийн шинжилгээ
- ХДШ-Хэвийн дискриминант шинжилгээ
- СЭЗДС- Санхүү Эдийн Засгийн Дээд Сургууль
- СЭЗИС- Санхүү Эдийн Засгийн Их Сургууль
- ШДШ- Шугаман дискриминант шинжилгээ

ОРШИЛ

Дискриминант шинжилгээ нь судлаачдад хамааран хувьсагч болон нэг эсвэл түүнээс дээш үл хамааран хувьсагчдын хоорондын уялдааг тодруулан өгдөг статистикийн арга хэрэгсэл юм. Дискриминант шинжилгээ нь регрессийн шинжилгээ болон вариацийн шинжилгээтэй төстэй боловч хамааран хувьсагчийн шинж чанараас хамааран ялгаатай байдаг. Дискриминант шинжилгээ нь судлаачийг хамааран хувьсагчийн хэмжүүрүүдтэй байх ба үл хамааран хувьсагчийг бүлэгт хуваагдсан байх шаардлагыг бий болгодог. Регресс болон вариацийн шинжилгээнд хамааран хувьсагч заавал “тоон өгөгдөл” байх шаардлагатай байдаг бол дискриминант шинжилгээнд хамааран хувьсагчийг “чанарын өгөгдөл” ба хэдэн ч бүлэгт хуваасан бай шинжилгээ хий боломжтой арга зүй юм.

Судлаачид энэхүү шинжилгээг хийхийн тулд хувьсагчаа хэрхэн сонгох, ялгах тайлбарлах гэх мэтээр асуудлуудтай тулгардаг. Түүнчлэн ангиллыг бий болгохын тулд нэг ба түүнээс олон математик тэгшитгэлүүдийг бий болгодог ба үүнийг дискриминант функц гэнэ. Бүлэг хоорондын ялгааг бий болгохын тулд ашиглагдах шинж чанаруудыг “ялгах хувьсагч”¹ гэж нэрлэдэг. Эдгээр хувьсагчид нь заавал интервал эсвэл хэмжүүрээ харьцаагаар илэрхийлсэн тохиолдолд вариацийг бодож, математик тэгшитгэлээр тайлбарлах боломжтой болдог. Ангиллын хувьсагчдын хувьд заавал заасан тоо байдаггүй учир нийт тохиолдлын тоо ихсэх тусам хувьсагчдын тоо хоёроос илүүгээр өсдөг байна. “Шугаман хослол” гэдэг нь нэг ба түүнээс олон хувьсагчдын ижил жинлэгдсэн шугаман хослолыг харуулдаг. Энэхүү шинжилгээг хийх үед шугаман хослол байх учиргүй байдаг. Хэрэв хоёр хувьсагчид төгс хамааралтай байвал тэр хоёр хувьсагчдыг нэгэн зэрэг оруулан үнэлгээ хийх нь боломжгүй байдаг. Өөр нэгэн нөхцөл бол олон судалгааны ажлуудад хүн амын ковариацийн матрицыг ашиглан бүлэг тус бүрт тэнцүү байх хэрэгтэй гэж үздэг. Иймээс хамгийн их хэрэглэгддэг арга бол “шугаман” дискриминант функц ба энэ нь энгийнээр шугаман хослолыг харуулах арга байдаг.

Энэхүү шинжилгээг 1960-аад онд бий болгож хөгжүүлэн дэвшүүлсээр ирсэн ч дэлхий даяар хийгдсэн судалгаанууд ихэвчлэн тухайн шинжилгээг илүү өргөн хүрээнд онолын хөгжүүлэлт хийсэн судалгаа түлхүү байгаа боловч эмпирик тохиолдолд хийсэн судалгаа цөөн тооны байна. Манай улсын хувьд сүүлийн жилүүдэд вариацийн шинжилгээ, логистик регрессийг ашиглан чанарын өгөгдөл дээр судалгааг хийж байгаа боловч энэхүү хараахан таньж, судалгаа хараахан хийгдээгүй байна. Тиймээс энэ төрлийн загваруудыг судалж, арга зүйн хэрэглээг нэмэгдүүлэх зорилготой байгаа юм.

Иймд судалгааны ажлын зорилгын хүрээнд дараах зорилтуудыг дэвшүүлж байна.

- Дискриминант шинжилгээний хэрэглээ: Ямар төрлийн хувьсагчдыг ашиглан шинжилгээ хийх боломжтой талаар судлах
- Дискриминант шинжилгээний онолын талын мэдлэгтэй болох хоёр ба түүнээс дээш бүлэгт хувьсагчдыг хэрхэн авч үзэх, нөхцөл ба таамаглалыг тодорхойлох

¹ ялгах хувьсагч- discrimination variable

- Дискриминант шинжилгээний арга зүйг тодорхойлох: хэрхэн, ямар программыг ашиглан хийж болох
- Дискриминант шинжилгээнээс гарсан үр дүнг тайлбарлаж, ядаж өөр ижил арга зүйтэй харьцуулах

Судалгааны ажлын хүрээнд ДШ-тэй холбоотой хийгдсэн судалгааны ажлын тойм, ДШ-ний арга зүй, эмпирик үнэлгээ гэсэн гурван бүлэгтэй. Эхний бүлэгт чиглэл чиглэлээр ангилан таван бүрдэл хэсэгт хуваан судалгааны ажил тус бүрийг зорилго, сонгож авсан хувьсагчид, шинжилгээ хийхдээ ашигласан асуулга зэргийг дурдан хураангуйллаа. Хоёрдугаар бүлэгт ДШ-ний арга зүйг дэлгэрэнгүй судалж, хоёр бүлэгтэй эсвэл олон бүлэгтэй үед хэрхэн тусгаарлах, ач холбогдлыг шалгах, дүрслэх зэргийг тус бүр харуулж өгөв. Харин сүүлийн бүлэгт энэхүү дипломын ажлын хүрээнд олж авсан онол арга зүйн мэдлэгийг ашиглан эмпирик судалгааны хэсэгт бодит өгөгдөл ашиглан оюутнуудын академик гүйцэтгэлийг харуулах шинжилгээг *Microsoft excel 2013* программын өгөгдлийн шинжилгээний хэрэгсэл “*Realstat*”-ыг ашиглан хийж үр дүнг харуулан, хавсралтанд дэлгэрэнгүй хавсаргав.

I БҮЛЭГ. ДИСКРИМИНАНТ ШИНЖИЛГЭЭНИЙ АРГА ЗҮЙН ТАЛААРХ СУДЛАГДСАН БАЙДЛЫН ТОЙМ

Дискриминант шинжилгээг ашиглан чанарын хамааран хувьсагч ба тоон үл хамааран хувьсагчдын хооронд шинжилгээ хийх боломжтой байдаг. Чанарын өгөгдөлтэй хамааран хувьсагчийг бүлгүүдэд хуваарилан шинжилгээ хийж болдгоороо бусад шинжилгээний аргуудаас давуу талтай. Энэхүү шинжилгээ нь вариацийн болон регрессийн шинжилгээтэй адил мэт боловч хамааран хувьсагчдыгаа шинж чанараас хамааран ялгаатай байдаг. Логистик регрессийн хувьд чанарын өгөгдлийг ашиглан бүлгүүдэд хуваарилан шинжилгээ хийх боломжтой байдаг ч, чанарын өгөгдлийг зөвхөн хоёр бүлэгт тусгаарладаг. Харин ДШ-ний хувьд чанарын өгөгдлийг хэдэн ч бүлэгт, хэдэн ч хамааран хувьсагчтайгаар шинжилгээ хийж ангилал хийж болон арга зүй юм.

1.1 Оюутнуудын амжилтыг дискриминант шинжилгээ ашиглан хийсэн судалгааны ажлууд

Загребийн Их Сургуулийн Байгууллага ба Мэдээллийн Факультетийн оюутнуудын сурлагын амжилтанд тулгуурлан 30 хувьсагчдын нөлөөлийг шинжилхэд ДШ-ийг ашигласан. Өгөгдлийг 2006/07 оны хичээлийн жилийн Мэдээлэл, Бизнесийн Системийн Бакалаврын ангийн 2-р курсийн оюутнууд болон 2008/09 оны Мэдээллийн Системийн Бакалаврын ангийн 3 болон 4-р курсын оюутнуудаас асуулга авч өгөдлөө бэлдсэн. Энэ судалгаа нь Байгууллага болон Мэдээллийн Факультетийн оюутнуудын сурлагын амжилтын урьдчилан таамаглагч хувьсагчдыг сонгоход чиглэгдсэн. Эхний асуулгад нийт 132 оюутнаас 110 хамрагдсан бол хоёр дох асуулгад 219 оюутнаас 113 оюутан хамрагдсан юм. Эхний асуулгыг хичээлийн хүрээнд явуулсан ба 34 хувьсагчийг авч үзсэнээс 31 нь сургалтын явцад хандах хандлагыг үнэлэх хувьсагчид байв. Асуулгын хариултыг тухайн нөхцөлийг хүлээн зөвшөөрөх эсвэл зөвшөөрөхгүй гэсэн хэлбэрээр авсан. Хувьсагчдыг хичээлийн бэрхшээл, хичээлийн агуулга, хичээлийн харилцаа, мэдээлэл технологийн хэрэгсэл болон уран зохиол, сургалтын арга, сургалтын дэмжлэг болон бусад гэсэн долоон категорт хуваасан. Харин хоёр дахь асуулгын хувьд 36 хувьсагчдыг хамааруулан үзсэн. Үүнд оюутны хүчин чармайлт (22), ерөнхий боловсролын сургууль (3), хүмүүжил (4), ололт, амжилтын үндэс (1), хувь хүний сургалтын хэв маяг (1), сурлагын амжилт (4) болон хүйс (1) зэргийг хамааруулсан. Нийт хувьсагчдын 18-д Likert scale-ийг ашигласан харин бусад 17 хувьсагчдад асуулт хариулт буюу заримд нь нээлттэй асуулга тавьсан юм. 30 хувьсагчдаас бүрдсэн шинжилгээний үр дүнгээс үзэхэд 8 нь амжилттай болон амжилтанд хүрээгүй оюутнуудыг хооронд нь ялгах нь ихээхэн хувь нэмэр оруулжээ. Үүнд элсэлтийн шалгалтын оноо, хувийн хариуцлага, суралцах хэв маяг, хичээл, анги доторх үйл ажиллагаан дахь цагийн хуваарилалт зэрэг хувьсагчид хамаарах юм (Блазенка Дивяк, Диеина Орески, 2009).

Гүржийн Технологийн Их Сургуулийн инженерийн ангийн оюутнууд физикийн танилцуулах хичээлд хангалттай үнэлгээ авч чаддаггүй ба үүнээс шалтгаалан эхний улиралд үздэг “Цахилгаан Соронзлол”-ын хичээл (дараагийн гурван улиралд үзэх хичээлүүдийн гол бүрэлдэхүүн хэсэг гэж үздэг)-д энэ сургуулийн бүх оюутнуудын 30

хүртэлх хувь нь хангалтгүй үнэлгээ (D-ээс F) авдаг байна. Иймээс хангалтгүй үнэлгээ авсан хичээлийг дахин судлах нь тухайн сургууль болон оюутнуудад ихээхэн зардал учруулдаг тухай таван жилийн хугацаанд “А” профессорын ангид хичээлийг нь судалж буй 1622 оюутнуудаас өгөгдлөө цуглуулан дискриминант шинжилгээг хийсэн. Энэ нь хичээлийн гүйцэтгэлийг урьдчилан таамаглах ба эрсдэлд байгаа оюутнуудыг тогтоох арга нь боловсролын аливаа оролцоог үнэлэх, төлөвлөх, хэрэгжүүлэхэд маш их ач холбогдолтой байдаг. Энэхүү кейс судалгаагаар инженерийн мэргэжлээр суралцагчдын хичээлийн гүйцэтгэлийн урьдчилан таамаглахын тулд дискриминант шинжилгээг ашигласан. Нийт 15 боломжит урьдчилан таамаглах хувьсагчийг (SAT, голч оноо г.м) судалж, аль хувьсагч нь эрсдэлд байгаа оюутнуудыг хамгийн сайн тодорхойлж, ач холбогдолтой байгааг тогтоосон байна. Дискриминант шинжилгээ нь амжилттай суралцагчдыг сайн тодорхойлж байгаа хэрнээ хангалтгүй суралцагчдыг зөвхөн 50 хувьтай таамаглаж чадаж байгаа юм. Оюутнуудын сурлагын бүртгэлээс харахад сайн суралцагчдын мэдээлэл нь байнгын бүртгэлтэй байгаа хэрнээ, муу суралцагчдын бүртгэл хангалтгүй байгааг дурдах хэрэгтэй. Мөн өөр аргатай харьцуулахад дискриминант шинжилгээ нь бусад шинжилгээг бодоход таамаглал, ач холбогдлыг илүү нарийн гаргаж байгаа бөгөөд өгөгдлийг маш үр дүнтэйгээр ашиглан, урьдчилан таамаглах боломжгүй хувьсагчдыг ч оруулан шинжилдэг нь давуу тал болж байгааг харуулж байв (Эдуард В. Томас, М. Жэксон Марр, Адриан Томас, Роберт М. Хүйм, 2014).

Нигерийн баруун өмнөд муж дахь их дээд сургуулиудад шинжлэх ухааны суурьтай Математикийн ангийн оюутнуудын гүйцэтгэлийг таамаглах зорилгоор дискриминант шинжилгээ хийсэн. Энэ ангид суралцагчдын хичээлдээ унах нь их байдаг (30%-аас илүү нь F дүн авдаг) байна. Шинжлэх ухааны хичээл нь суурь учир заавал суралцах шаардлагатай байдаг учир энэхүү хичээл дээр унах нь их сургуульд ч, оюутанд ч тодорхой хэмжээний зардал учруулдаг. Сургалтын хөтөлбөрийг урьдчилан таамаглах, урьдчилж таамаглах хүчин зүйлсийг тодорхойлж болох арга барилыг тодорхойлж, эрсдэлд байгаа оюутнуудыг мэдсэнээр боловсролын хөтөлбөрийг төлөвлөх, хэрэгжүүлэх, үнэлэхэд ач холбогдолтой юм. Өгөгдөл нь 1600 шинжлэх ухаан суралцагчдын оноон дээр дүн шинжилгээн дээр тулгуурласан юм. Нийт 15 боломжит таамаглалын хувьсагчийг тодорхойлж, эрсдэлд байгаа оюутнуудыг аль хувьсагч хамгийн сайн илэрхийлж байгааг тодорхойлох нь чухал юм. Оюутнуудын бүртгэлээс авах боломжит мэдээллийг ашиглан оюутнуудын 50% нь хичээлдээ унасан гэж ангилж чадсан (Абаёоми А. Акомлафе, Жи Эн. Амахиа, 2015).

“Бизнесийн удирдлагын магистр” (МВА) чиглэлээр суралцагчдын академик гүйцэтгэлд нөлөөлөгч хүчин зүйлсийг харгалзан ДШ хийсэн. Шаардлагатай өгөгдлүүдийг цуглуулахдаа “Uttaranchal их сургууль”-ийн нийт 102 МВА суралцагч оюутнуудаас цуглуулж холбогдох бүх мэдээллийг сургуулийн архиваас авч энэхүү судалгаанд дүн шинжилгээ хийсэн. Магистрын зэргээр суралцаж буй суралцагчдын академик гүйцэтгэлийг хамааран хувьсагчаар авч харин оюутны ирц, түүний нийгэмд эзлэх байр суурь, хичээлийн явц, анги хоорондын ялгаа, элсэлтийн шалгалтын дүн зэргийг үл хамаарах хувьсагчдаар авч үзсэн. Ийнхүү үзэхдээ оюутнуудыг амжилттай болон амжилтгүй гэдгээр хувааж чадах чухал хувьсагчдын нөлөөг авч үзэхэд ирц, сургалтын явц чухал нөлөөтэй байсан юм. Ийнхүү эдгээр хувьсагчдыг ашиглан ангилал хийхэд

зөвхөн нэг хувьсагч “нийгэмд эзлэх байр суурь” нь сөрөг коэффициенттэй гарч дискриминант функцэд сөрөг нөлөө үзүүлсэн. Ийнхүү ангиллыг зөв хийсэн эсэхийг шалгаж үзэхэд 81.4% зөв ангилж чадсан байсан юм (Ди.Эс.Чаурибел, Хэм Чандра Котари, Сумати Карпүр, 2016).

Олон улсын үнэлгээгээр үл хамаарах хувьсагчид болох докторт суралцагч оюутнуудын харьцаа (2012-2013 оны хичээлийн жилийн докторын ангийн оюутнууд/тус оны суралцагч нийт оюутнууд) , багш нарт ногдох нэг оюутны тоо (2012-2013 оны хичээлийн жилийн нийт оюутнуудын тоо/2012 оны багш нарын тоо) мөн нийтлэлийн оноо 2013 оны SCI, SSSC, AHCI шинжилгээнд хамрагдсан нийтлэлийн тоо) зэргийг харгалзан Туркийн их дээд сургуулиудын зэрэглэлийг нарийвчлан тогтоох зорилготой судалгаа юм. Судалгааны өгөгдлийг Туркийн Их Сургуулийн Зэрэглэлийг Академик Гүйцэтгэлээр гаргасан. Өгөгдлийг олон улсын их дээд сургуулиудын эхний 500-д багтдаг (бүлэг А) , 500-1000-д багтдаг (бүлэг Б) , 1000-1500-д багтдаг (бүлэг В) гэж гурван бүлэгт хуваасан. Дискриминант шинжилгээ нь өгөгдөлд дүн шинжилгээ хийхэд ашиглагдав. Судалгааны үр дүнгээс харахад “докторын хөтөлбөрт суралцагч оюутнуудын харьцаа” бүлгүүдийг ялгахад хамгийн их хувь нэмэр оруулсан хувьсагч юм. Ангиллын үр дүнд А бүлгийн 10 их сургуулиас 9 (90%), Б бүлгийн 15 их сургуулиас 10 (66.7%), В бүлгийн 26 их сургуулиас 19 (73.1%) нь зөв ангилагдсан байна (Метин Озкан, 2017).

1.2 Зах зээл, хөгжлийн асуудлуудын мэдээллийг ашиглан олон улсын дискриминант шинжилгээ хийсэн судалгааны ажлууд

Европийн статистикийн үндэсний газраас авсан Румын улсын 42 мужийн түвшин дэх хөгжлийн багц дээр явуулсан ба хөгжлийн найман бүсэд хуваан статистикийн ажиглалтыг хийсэн. Энэ улсын Бухарест мужийн өгөгдөл нь гажуудалтай учир энэхүү шинжилгээнд оруулж үнэлээгүй болно. Энэхүү өгөгдлийг 2003 онд цуглуулсан бөгөөд тухайн бүс нутгийн боломжит статистикийн мэдээлэл байсан юм. Тус судалгааны үл хамаарах хувьсагчид нь Румыний эдийн засаг нийгмийн хөгжлийг тодорхойлох хувьсагчид байна. Эдгээрт *phys2003* (1000 хүнд ногдох эмч нарын тоо), *urb2003* (хотжилтын түвшин), *wage2003* (нэрлэсэн цэвэр дундаж цалин), *stud2003* (дээд боловсролын сургуулийн хүн ам), *work2003* (ажлын байрандаа осолдсон хүмүүсийн тоо), *RD2003* (судалгаа ба хөгжлийн зардал), *pop2003* (оршин суугчдын тоо), *libr2003* (номын сангийн тоо) багтана. Мужуудийг бүлэг болгон хуваахдаа хөдөлмөрийн бүтээмжийн хувьд маш багаас (1), маш их (4) хүртэл авч үзсэн. Хамгийн эхлээд найман хувьсагчдаасаа ялгаварлах² хувьсагчдын сонгосон бөгөөд эдгээрээс *work2003*, *pop2003*, *libr2003* зэрэг нь ач холбогдолгүй гасан тул хасах хэрэгтэй гэсэн дүгнэлттэй тулгарсан. Мужуудыг 4 бүлэгт хувааснаас өөрийн утгуудыг тооцоод үзэхэд хамгийн их гурван хувийн утгуудтай дискриминант функцүүд бий болсон. Эхний дискриминант функцийн хувьд эхний гурван хувьсагчид нь хамгийн их утгатай буюу хамааралтай байсан, хоёр дахь дискриминант функцийн хувьд дараагийн дөрвөн хувьсагчид хүчтэй хамааралтай байгааг харуулсан бол гурав дахь дискриминант фунцын хувьд хамгийн сүүлийн хувьсагчтай хамааралтай байгаа юм. Дискриминант функц дээр үндэслэн мужуудын 65.9

² ялгаварлах- discriminating

хувь нь үнэ зөв ангилагдсан гэж үзэж байгаа бөгөөд дискриминант шинжилгээгээр улсын хагжлийн үе шатыг хөдөлмөрийн бүтээмжээр ангилан үнэлэхэд хангалттай гэж үзсэн (Элизабета Жаба, Донат Вазил Жемна, Даниала Виорика, Кристиана Брижжет Балан, 2007).

Гана улсын хойд хэсэг болох Хойд Ганад хамгийн их ядууралтай газар гэж тодорхойлогддог. Санхүүгийн оролцоо, ядуурлыг бууруулах хоорондын уялдаа холбоог харгалзан үзээд, энэ судалгааны зорилго нь дискриминант функцийн загварыг ашиглан Хойд Гана дахь санхүүгийн оролцоотой эрэлт хэрэгцээний саад бэрхшээлийг тодорхойлоход оршино. Энэ судалгаанд хамаарах хувьсагчаа санхүүгийн оролцоо хэмээн авсан бөгөөд энэ нь нийлүүлэлтийн талын шалгуур үзүүлэлтүүдийг ашиглан тооцдог байна. Гэвч энэхүү судалгаанд “санхүүгийн бүрэн оролцоотой улс нь тэдгээрийг ашиглаж чадах хүмүүст боломжийн үнийг санал болгож, үйлчлүүлэгчдээ нэр хүнд бүхий чанартай үйлчилгээг хүргэх боломжтой” гэсэн тодорхойлолт дээр үндэслэн эрэлтийн талын индикаторуудыг ашиглан санхүүгийн оролцоог хэмжих болсон. Нэг өрх дор хаяж дараах үзүүлэлтүүдийг нэг санхүүгийн үйлчилгээг авсан байх хэрэгтэй гэж үзсэн. Үүнд: албан ёсны банкны данс, тэтгэвэрийн сан, ипотекийн зээл, кредит/дебит карт, даатгал, бичил санхүүгийн зээл, гар утасны төлбөрийн данс, бонд, хувьцаа зэрэг хамрагдсан. Санхүүгийн оролцоо нь энэхүү судалгаанд бинар хувьсагч бөгөөд дээрх нөхцөлүүдийн аль нэгийг өрх хангаж байвал 1 үгүй бол 0 гэсэн утгыг авсан. Харин үл хамаарах хувьсагчид нь хандлага (зарцуулалтын дадал зуршил), итгэл (санхүүгийн институци нь таны зүрхэнд байх), зай (санхүүгийн байгууллага үйлчилгээтэй ойрхон), зардал (санхүүгийн үйлчилгээний хүртээмж), соёл (соёл болон зан заншлын итгэл үнэмшил), санхүүгийн чадавх (мэдлэг, ур чадвар, хүртээмжинд суурилсан санхүүгийн нөөцийг үр дүнтэй удирдах чадвар), баримт бичиг, мөнгө, бичиг үсэг тайлагдсан байдал. Үр дүнд хамаарах хувьсагч болох санхүүгийн оролцооны хэмжилтийг нэмэгдүүлсэн ч нийт асуулганд хариулсан 378 хүнээс 169 (44.7%) нь санхүүгийн хувьд хамрагдсан гэсэн бол үлдсэн 209 (55.3%) нь санхүүгийн хувьд хамрагдаагүй байна. Энэ нь хөгжиж буй улс орнууд (Нигери, Уганда, Пакистан, Гана, Египет, Здагч ертөнцийн орнууд) санхүүгийн индикаторуудын ихэнх үзүүлэлтүүдээс хол хоцорсон гэсэн мэдээлэлтэй нийцтэй байгаа юм. Мөн Гана улсын 1000 хүн тутмын 40% нь л санхүүгийн зах зээлд ямар нэгэн байдлаар хамрагдсан байдаг бол үлдсэн 60% нь оролцоогүй байна. Түүнчлэн хийгдсэн судалгаагаар нийт насанд хүрэгчдийн 29% нь албан ёсны банкны данстай байдаг. Мөн санхүүжилтын оролцоотой эрэлт хэрэгцээний саад бэрхшээлүүдээс хамгийн эхэнд “соёл” ордог гэж гарсан юм. Учир нь Хойд Ганагийн зонхилох шашин нь Ислам бөгөөд сонирхлын зөрчилд автан банкнаас зээл авах болон санхүүгийн бусад үйлчилгээг авахаас татгалзсан ойлгомжтой юм. Зардал, чадавх, итгэл, бичиг үсэг тайлагдсан байдал, зай, хандлага зэрэг нь санхүүгийн оролцоонд ноцтой саад бэрхшээл болж байна гэж үзсэн. Гэхдээ мөнгө, баримтжуулалт, гэр бүл зэрэг нь харин санхүүжилтэнд шаардагдах эрэлт хэрэгцээний саад бэрхшээл болохгүй байгаа юм. Тиймээс санхүүгийн оролцоотой эрэлтийн талын саад бэрхшээлийг бууруулахын тулд засгийн газрын бодлого нь ядуурлыг бууруулахад хүргэж буй бүс нутагт санхүүгийн оролцоог нэмэгдүүлнэ гэж үзсэн (Иссахаку Якуби, Романус Динез, Даниел Буор, Вахаб А. Идрусу, 2017).

1.3 Банкны мэдээллийг ашиглан дискриминант шинжилгээ хийсэн ажлууд

Хоёр бүлгийн ангиллын асуудлыг шийдвэрлэхийн тулд өгөгдлийн бүрхүүлийн шинжилгээ гэсэн бенчмаркинг загварыг ашиглахад энэ судалгааны зорилго чиглэгдсэн. Энэхүү санаа нь бенчмаркууд өөрсдийн үйлдвэрлэлийн боломжит багцуудад нөлөөлдөг гэснээс үүдэлтэй. Хэрэв ажиглалтууд ижил бүлэгт байвал тэд ижил үйлдвэрлэлийн боломжит багцад байх ба ижил бенчмаркуудаар нөлөөлдөг гэсэн үг байдаг. Түүнчлэн хоёр бүлгийн бенчмаркууд нь хос шугаман бус дискриминант хилийг тогтоохдоо бусад параметерт дискриминант аргуудын ангиллын функцийг урьдчилан таамаглахгүйгээр тодорхойлдог юм. Дискриминант хил нь хоёр бүлгийн хилээр тогтоогдсон тохиолдолд таамагласан процедур нь зарим дискриминант аргууд болох бүх ажиглалтуудыг бүлгийн дунджаас хамаарах хамаарлыг дискриминант функцийг бий болгохоос хамаарах юм. Судалгааны хүрээнд бодит өгөгдлийн багцыг санал болгосон хандлагын гүйцэтгэлийг (Benchmark*DA) шугаман дискриминант функц болон MSD аргуудаар тус тус үнэлсэн. Өгөгдлийн багцыг дампуурсан болон дампуураагүй пүүсүүдийн хувьд сонгосон. Хувьсагчдыг сонгохдоо ажиллах хүч/нийт хөрөнгө, хуримтлагдсан ашиг/нийт хөрөнгө, татварын өмнөх ашиг/нийт хөрөнгө, нийт өр төлбөрийн өөрийн хөрөнгийн үнэ цэн/нийт өр төлбөрийн зах зээлийн үнэ цэн, борлуулалт/нийт хөрөнгө гэсэн таван харьцааг авч үзсэн. Эдгээрт бүлгүүдийн давхцал нь буруу ангилагдах эх үүсвэр байх ба энэ асуудлыг шийдэхийн тулд хамгийн бага давхацсан хилийг олж шугаман хувиргалтыг ашиглах хэрэгтэй байдаг. Дээрх үнэлгээнээс гарсан үр дүн нь өгөгдөлд шугаман хувиргалт хийх нь ангиллын түвшинг сайжруулж болохыг харуулсан. Мөн өгөгдлийн багцын ажиглалтын тоо ижил биш, гажуудал ихтэй байх нь бүлэг томрох тусам гажуудал их илрэхийг харуулж байна. Гэсэн хэдий ч хандлагын гүйцэтгэлийн арга болох Benchmark*DA нь бусад аргуудтай харьцуулахад жижиг бүлгүүдийг ангилах нарийвлал сайтай байгаа бөгөөд өгөгдлийн олонлогт гажуудал ихтэй байх нь хоёр үйлдвэрлэлийн боломжит багцад том огтлолцол үүсгэсэн юм (Дон Сан Чан, И Чун Куо, 2007).

Банкны ашигт ажиллагааны түвшний гол онцлогуудыг тодруулах, тайлбарлахын тулд дискриминант шинжилгээг энэхүү судалгааны хүрээнд хийсэн. Түүврээ Хорват улсын 2003 (41 банкны түүвр), 2008 (33 банкны түүвр) онуудад идэвхтэй ажиллаж байсан банкуудын балансын ерөнхий утгууд болон орлогын мэдээллүүдээс сонгосон. 2003 оны хувьд 1990 ээд оны банкны системийн хямрал ба 2002 оны дунд зэргийн бакуудын дампуурлын дараа банкны тогтвортой байдлыг харуулсан эхний жилийг илэрхийлнэ. Харин 2008 оноос хойш улс орны бүтцийн асуудлуудад үндэслэн эдийн засгийн тогтворгүй байдал улам даамжирсан учраас энэхүү онуудыг сонгосон. Хамааран хувьсагчдаар чанарын өгөгдөл харин үл хамааран хувьсагчдаар тоон өгөгдөл авсан. Ашиг, алдагдлын бүртгэлийг тийм эсвэл үгүй (1,0), харин дундаж өмчийн өгөөжийг дундаж хэмжээнээс давсан тохиолдолд 1, эсрэг тохиолдолд 0 гэж тэмдэглэсэн. Өгөгдлөөс үзэхэд өр төлбөр болон хөрөнгийн өсөлт нь банкны ашигт ажиллагаанд эерэг хамааралтай байна. Харин өмчийн санхүүжилт өндөр байх нь ашигт ажиллагаатай сөрөг хамааралтай байгаа хэдий ч өмчийн дундаж өгөөжөөс дээгүүр эерэг хамааралтайг шинжлэн харуулсан. Өндөр төлбөрийн орлоготой байх нь ашигт ажиллагаанд эерэг нөлөө үзүүлдэг байна. Учир нь банкны өрсөлдөөнт салбарт цэвэр хүүгийн хэмжээ байнга буурч байгаа үед хадгаламжийн үйлчилгээг санал болгодог банкуудаас илүү төрөл

бүрийн төлбөрийн үйлчилгээг санал болгодог банкууд давуу талтай байдаг. Түүнчлэн банкууд бүтээгдэхүүний төрлөө өсгөвөл өрсөлдөх чадвар сайтай болдог. Банкны ашигт ажиллагааг тогтвортой байлгах үүднээс авто-санхүүжилтийн үйл явц, тогтвортой байдлын бүтцийг бий болгохын зэрэгцээ банкны ашиг орлогыг судлах нь нэн чухал асуудал хэвээр байгаа билээ. Энэ судалгаанд дискриминант шинжилгээ хийж дүгнэснээр олон хүчин зүйлийн статистик аргыг ашиглах замаар тэдгээрийн санхүүгийн харьцааг банкуудын ашигт болон ашигт бус бүлгүүдийн ангилж, дундаж ашгийн түвшинг өндөр нарийвчлалтайгаар авч үзсэн юм (Роза Анте, Кундид Ана, 2013).

1.4 Хэрэглэгчийн худалдан авалттай холбоотой дискриминант шинжилгээг ашиглан хийсэн судалгааны ажлууд

Энэтхэгийн жижиглэн худалдааны салбар нь өрсөлдөх чадвартай жижиглэн худалдааны брендүүдтэй тулгарч байгаа бөгөөд хэрэглэгчид сонголтын хувьд илүү хуваагдмал болсон. Маркетингийн орчин нь жижиглэнгийн худалдаачдад илүү ярвигтай болсон. Энэ утгаараа энэхүү судалгаагаар зарим тогтсон таамаглалуудын суурь болон сегментийг сэжигтэй болон хэтийн төлөв гэж ялгасны үндсэн дээр дискриминант тэшитгэлийг томьёолохыг оролдсон. Судалгаа нь брендийн мэдээлэл хангалттай түвшинд байгаа тохиолдолд зорилтот сегментийн зорилгыг урьдчилан таамаглахыг зорисон. Жижиглэн худалдаа эрхлэгчид нь энэхүү аргыг ашиглан өөрийн брендийн талаарх зах зээлийн үнэлгээ болон зах зээлийн эрсдэлийг бууруулахад тусалдаг. Худалдан авалт хийхэд хадалдан авагчийн санаа бодол эсвэл сэжигтэй байдал шалтгаалдаг эсэхийг урьдчилан таамаглахын тулд дискриминант шинжилгээ хийсэн. Урьдчилан таамаглагч хувьсагчид нь нас, Cherokee мөнгөний үнэ цэн, Cherokee –ийн үндсэн бүтээгдэхүүний шинж чанар, Cherokee брендийн онцлог зэрэг багтсан бөгөөд ажиглагдсан урьдчилан таамаглагч бүрийн хувьд ач холбогдлын дундаж өөр байв. Шалгуур үзүүлэлтүүдийг тодорхойлоход нэлээд төстэй байсан бөгөөд Vox's M нь ковариацийн тэгшитгэлийн таамаглалыг хүлээн зөвшөөрсөн байна. Дискриминант функц нь бүлэг хоорондын болон бүх урьдчилан таамаглагчдын хоорондын ач холбогдолтой эсэхийг харуулсан. Бүлэг хоорондын хувьсагчид 72.6%-ийг эзэлж байгаа боловч бүтцийн матрицийн ойролцоо дүн шинжилгээ хийсэн нь зөвхөн хоёр ач холбогдолтой таамаглагч болох Cherokee “мөнгөний үнэ цэн” бренд (0.790), Брендийн нэр хүндийн Cherokee оноо (0.513) нарыг нас, бүтээгдэхүүний шинж чанарыг харгалзан үзэхэд муу таамаглагчид гэж үзсэн. Хэрэглэгчийн худалдан авах зорилгын хувьд брендийн зах зээлийн бодит байр суурийг үнэлэх зорилгоор энэхүү судалгааг зах зээлд ашигладаг (Сугата Банержи, Сарват Павар, 2014).

1.5 Дискриминант шинжилгээний хөгжүүлэлттэй холбоотойгоор хийгдсэн судалгааны ажлууд

ИХБТХӨ гэдэг нь түүврийн хэмжээнээс өгөгдлийн векторын хэмжээс нь их байхыг хэлдэг. ИХБТХӨ-ийн өгөгдөл нь ихэвчлэн генийн илэрхийлэл микро иж бүрдлийн шинжилгээ, эмнэлгийн зургийн шинжилгээ, текстийн ангилал, нүүр царайг таних зэрэг байдаг. (Маррон, 2007) “өгөгдлийн овоолго” хэмээх ухагдахууныг ИХБТХӨ-ын тохиргоонд илрүүлсэн ба үүний утга нь өгөгдлүүд нь бие биенийхээ дээр байрладаг гэсэн

үг юм. Энэ нь ялган ангилхад тустай биш байдаг учир нь векторын харгалзах чиглэл нь туршилтын өгөгдлийн бодит байдлыг тодорхойлсон байдаг. Өгөгдлийн овоолох чиглэл нь бүлгүүдийн түүврийг тусгаарлахад маш сайн тохирдог гэвч энэ нь зайлшгүй байх ёстой түгээмэл бус шинж чанартай байдаг. Хэрэв цөөн тооны томоохон хувьсагчид нь ялган шинжилхэд хангалттай байвал эдгээр хувьсагчдыг тодорхойлох нь туршилтын өгөгдөлд “өгөгдлийн овоолго”-оос урьдчилан сэргийлэх, улмаар түүврийн сайн ангиллыг гаргахад тусалдаг юм. Эдгээрийн үр дүнд сийрэг дискриминант вектор байгуулахад зориулсан шинэ алгоритмыг санал болгосон. Сийрэг дискриминант вектор нь их хэмжээст өгөгдлийн хувьд ажиглалтын хэмжээсийг бууруулахад ашигтай байдаг. Сонгодог Фишерийн шугаман дискриминант шинжилгээний хэрэглээ нь их хэмжээст, түүврийн хэмжээ багатай байдаг тул өгөгдөл үүсгэх асуудал үүсдэг. Дискриминант векторуудад сийрэгжилтийг нэвтрүүлэх нь өгөгдлийг овоолох, түүнтэй холбоотой асуудлыг арилгахад үр дүн ихтэй байдаг. Судалгаанд үнэлсэн загварчилсан болон бодит өгөгдлийн жишээнүүд нь үл хамаарах хувьсагчууд байгаа үед сийрэг шугаман дискриминант шинжилгээний аргаар чухал хувьсагчдыг сонгож болох ба ингэснээр ангиллын талбар сайжрах болно (Зихуа Чиао, Лан Зхоу, Жианхуа З.Хуанг, 2007).

II БҮЛЭГ. ДИСКРИМИНАНТ ШИНЖИЛГЭЭНИЙ ОНОЛ, АРГА ЗҮЙ

Дискриминант шинжилгээ нь судлаачдад хамааран хувьсагч болон нэг эсвэл түүнээс дээш үл хамааран хувьсагчдын хоорондын уялдааг тодруулан өгдөг статистикийн арга хэрэгсэл юм. Дискриминант шинжилгээ нь регрессийн шинжилгээ болон вариацийн шинжилгээтэй төстэй боловч хамааран хувьсагчийн шинж чанараас хамааран ялгаатай байдаг. Дискриминант шинжилгээ нь судлаачийг хамааран хувьсагчийн хэмжүүрүүдтэй байх ба үл хамааран хувьсагчийг бүлэгт хуваагдсан байх шаардлагыг бий болгодог. Регресс болон вариацийн шинжилгээнд хамааран хувьсагч заавал “тоон өгөгдөл” байх шаардлагатай байдаг бол дискриминант шинжилгээнд хамааран хувьсагчийг “чанарын өгөгдөл” ба хэдэн ч бүлэгт хуваасан бай шинжилгээ хийх боломжтой арга зүй юм.

Судлаачид энэхүү шинжилгээг хийхийн тулд хувьсагчаа хэрхэн сонгох, ялгах тайлбарлах гэх мэтээр асуудлуудтай тулгардаг. Түүнчлэн ангиллыг бий болгохын тулд нэг ба түүнээс олон математик тэгшитгэлүүдийг бий болгодог ба үүнийг дискриминант функц гэнэ. Бүлэг хоорондын ялгааг бий болгохын тулд ашиглагдах шинж чанаруудыг “ялгах хувьсагч” гэж нэрлэдэг. Эдгээр хувьсагчид нь заавал интервал эсвэл хэмжүүрээ харьцаагаар илэрхийлсэн тохиолдолд вариацийг бодож, математик тэгшитгэлээр тайлбарлах боломжтой болдог. Ангиллын хувьсагчдын хувьд заавал заасан тоо байдаггүй учир нийт тохиолдлын тоо ихсэх тусам хувьсагчдын тоо хочроос илүүгээр өсдөг.

“Шугаман хослол гэдэг нь нэг ба түүнээс олон хувьсагчдын ижил жинлэгдсэн шугаман хослолыг харуулдаг. Энэхүү шинжилгээг хийх үед шугаман хослол байх учиргүй байдаг. Хэрэв хоёр хувьсагчид төгс хамааралтай байвал тэр хоёр хувьсагчдыг нэгэн зэрэг оруулан үнэлгээ хийх нь боломжгүй байдаг.

Өөр нэгэн нөхцөл бол олон судалгааны ажлуудад хүн амын ковариацийн матрицыг ашиглан бүлэг тус бүрт тэнцүү байх хэрэгтэй гэж үздэг. Иймээс хамгийн их хэрэглэгддэг арга бол “шугаман” дискриминант функц ба энэ нь энгийнээр шугаман хослолыг харуулах арга байдаг.

Дискриминант шинжилгээг хийхдээ бүлгүүдийг хооронд нь тусгаарлаж, хоёр бүлэгтэй тохиолдол болон түүнээс дээш бүлэгтэй тохиолдолд өөр өөр аргыг ашиглан хийдэг.

2.1 Бүлгийг тусгаарлах нь

Эх олонлогоос өөрөөс нь эсвэл эх олонлогоос түүвэр үүсгэж бүлгүүдийг бий болгодог. Бүлгүүдийг тусгаарлахад хоёр үндсэн нөхцөл байдаг.

1. Шугаман функцийн хувьсагчид нь хоёр болон түүнээс дээш бүлгүүдийн ялгааг дүрслэх эсвэл таниулахад бүлгийг тусгаарладаг. Тодорхойлогч дискриминант шинжилгээний зорилго нь тусгаарласан бүлгүүдийн хувьсагчдын харьцангуй хамаарлыг тодорхойлох мөн ямар цэгүүд бүлгүүдийн тохироог хамгийн сайн, оновчтой харуулж байгааг олдог.

2. Бүлгүүдийн таамаглал болон ажиглалтын хуваарилалтаас харахад шугаман болон квадрат функцүүдийн хувьсагчид нь оноогдсон түүврийн нэгжийг аль нэг бүлэг рүү хуваарилдаг.

2.1.1 Хоёр бүлгийн дискриминант функц

Хоёр эх олонлогийг ижил ковариацийн матрицтай харин ялгаатай дундаж вектор μ_1 болон μ_2 -той гэж үзнэ. Хоёр эх олонлогийн түүврүүд болох $y_{11}, y_{12}, \dots, y_{1n_1}$ ба $y_{21}, y_{22}, \dots, y_{2n_2}$ -тай ажилладаг. Ихэнхдээ y_{ij} -ийн вектор тус бүр нь p хувьсагчийн хэмжигдэхүүнээс бүрддэг. Дискриминант функц нь хоёр бүлгийн дундаж векторын зайг хамгийн их байлгах хувьсагчдын шугаман хослолыг хэлдэг. Шугаман хослол $z = \mathbf{a}'\mathbf{y}$ ажиглалтын вектор бүрийг скаляр болгон дараах байдлаар хөрвүүлдэг.

$$z_{1i} = \mathbf{a}'\mathbf{y}_{1i} = a_1 y_{1i1} + a_2 y_{1i2} + \dots + a_p y_{1ip}, \quad i = 1, 2, \dots, n_1$$

$$z_{2i} = \mathbf{a}'\mathbf{y}_{2i} = a_1 y_{2i1} + a_2 y_{2i2} + \dots + a_p y_{2ip}, \quad i = 1, 2, \dots, n_2$$

Тиймээс хоёр түүвэр дэх $n_1 + n_2$ таамаглалын вектор,

$$\begin{pmatrix} y_{11} & y_{21} \\ y_{12} & y_{22} \\ \vdots & \vdots \\ y_{1n_1} & y_{2n_2} \end{pmatrix},$$

нар нь скаляр болон хувирахдаа дараах хэлбэртэй болдог,

$$\begin{pmatrix} z_{11} & z_{21} \\ z_{12} & z_{22} \\ \vdots & \vdots \\ z_{1n_1} & z_{2n_2} \end{pmatrix}.$$

Дундажийг олохдоо $\bar{z}_1 = \sum_{i=1}^{n_1} z_{1i}/n_1 = \mathbf{a}'\bar{y}_1$ ба $\bar{z}_2 = \mathbf{a}'\bar{y}_2$ гэх ба үүнд $\bar{y}_1 = \sum_{i=1}^{n_1} y_{1i}/n_1$, $\bar{y}_2 = \sum_{i=1}^{n_2} y_{2i}/n_2$ байна. Вектор \mathbf{a} нь $(\bar{z}_1 - \bar{z}_2)/s_z$ гэсэн стандартчилагдсан зөрүүг хамгийн их байлгах хэлбэр юм. Хэрэв $(\bar{z}_1 - \bar{z}_2)/s_z$ сөрөг байж болох учраас квадратлаг хэлбэрийг илүү ашигладаг.

$$\frac{(\bar{z}_1 - \bar{z}_2)^2}{s_z^2} = \frac{[\mathbf{a}'(\bar{y}_1 - \bar{y}_2)]^2}{\mathbf{a}'S_{pl}\mathbf{a}} \quad (2.1)$$

(2.1) тэгшитгэл нь хамгийн ихдээ

$$\mathbf{a} = S_{pl}^{-1}(\bar{y}_1 - \bar{y}_2), \quad (2.2)$$

байх эсвэл \mathbf{a} нь ямар ч $S_{pl}^{-1}(\bar{y}_1 - \bar{y}_2)$ -ийн үржвэр байдаг. \mathbf{a} векторыг хамгийн их байлгах нь нэг байдаггүй. Гэхдээ энэ нь ганц чиглэлтэй буюу харьцангуй утга эсвэл a_1, a_2, \dots, a_p харьцаа ганц байна. S_{pl}^{-1} оршин байхын тулд $n_1 + n_2 - 2 > p$ байх хэрэгтэй.

\bar{y}_1, \bar{y}_2 холбогдох шугамтай $\mathbf{a} = S_{pl}^{-1}(\bar{y}_1 - \bar{y}_2)$ гэсэн оновчтой чиглэл нь үр дүнтэй паралель байх ёстой. Учир нь квадрат зай (distance) $(\bar{z}_1 - \bar{z}_2)^2/s_z^2$ нь \bar{y}_1, \bar{y}_2 хоёрын хоорондох стандартчилагдсан зайтай эквалент байна. Энэ нь (2.1) тэгшитгэл рүү (2.2) тэгшитгэлийг нэгтгэвэл

$$\frac{(\bar{z}_1 - \bar{z}_2)^2}{s_z^2} = (\bar{y}_1 - \bar{y}_2)' S_{pl}^{-1} (\bar{y}_1 - \bar{y}_2) \quad (2.3)$$

гэж илэрхийлэгдэж болно.

2.1.2 Хоёр бүлгийн дискриминант шинжилгээ ба олон хүчин зүйлсийн регрессийн хамаарал

Хамааран хувьсагч болон үл хамааран хувьсагч нь дискриминант шинжилгээ болон олон хүчин зүйлсийн регрессд урвуу байдаг. Дискриминант шинжилгээнд авч үздэг хамааран хувьсагч (y 's) нь регрессийн утгад үл хамаарах хувьсагч болдог.

w -г бүлгийн хувьсагч (бүлэг 1 болон 2-г хамааруулсан) гэж үзэн $\bar{w} = 0$ мөн y 's-д w тохирч байх үед регрессийн коэффициентийн вектор $\mathbf{b} = (b_1, b_2, \dots, b_p)'$ гэж тодорхойлогдоно. Үүний дараагаар \mathbf{b} -г дискриминант функцийн коэффициентийн вектор $\mathbf{a} = S_{pl}^{-1}(\bar{y}_1 - \bar{y}_2)$ -тай пропорциональ гэж үзэн

$$\mathbf{b} = \frac{n_1 n_2}{(n_1 + n_2)(n_1 + n_2 - 2 + T^2)} \mathbf{a}, \quad (2.4)$$

байх ба үүнд $T^2 = [n_1 n_2 / (n_1 + n_2)] (\bar{y}_1 - \bar{y}_2)' S_{pl}^{-1} (\bar{y}_1 - \bar{y}_2)$ байх ба олон хүчин зүйлийн корреляцийн коэффициент R^2 нь T^2 -тай холбогдохдоо дараах хэлбэрт бичигдэнэ.

$$R^2 = (\bar{y}_1 - \bar{y}_2)' \mathbf{b} = \frac{T^2}{n_1 + n_2 - 2 + T^2}. \quad (2.5)$$

$$F = \frac{n_1 + n_2 - p - q - 1}{q} \frac{R_{p+q}^2 - R_p^2}{1 - R_{p+q}^2}, \quad (2.6)$$

2.2 Олон бүлгийн дискриминант шинжилгээ

2.2.1 Дискриминант функцүүд

Олон бүлэгт дискриминант шинжилгээ хийх нь олон хувьсагчидтай k бүлгүүдийг хамгийн сайн тусгаарлаж чадах хувьсагчдыг олохтой холбогддог. Олон бүлгийн дискриминант шинжилгээ нь янз бүрийн зорилгоор ашиглагдах боломжтой байдаг.

1. Бүлгүүдийн тусгаарлалтыг хоёр хэмжээст талбайд дүрслэн шалгана. Хэрэв хоёроос олон бүлэг байвал бүлгүүдийн тусгаарлалыг дүрслэхдээ нэгээс олон дискриминант функц шаардагддаг. p хэмжээст талбай дээр дүрслэгдсэн цэгүүдийг

хоёр хэмжээст талбайд дүрслэхдээ эхний хоёр дискриминант функцийг ашиглана. Бүлгүүдийг тусгаарлахад аль хэлбэр нь хамгийн сайн байгааг таамаглана.

2. Бүлгүүдэд хуваасны дараах анхны олонлогийг харуулж чадахаар анхны хувьсагчдын дэд олонлогийг олно.
3. Бүлгүүдийн тусгаарлалтанд харьцангуй хамааралтай хувьсагчдыг олж эрэмбэлэнэ.
4. Дискриминант функцүүдээр илэрхийлэгдсэн шинэ хэмжээсүүдийг харуулна.

i -р бүлгийн n_i ажиглалтын утгатай k бүлэгт y_{ij} ажиглалтын вектор бүрийг $z_{ij} = a'y_{ij}$, $i = 1, 2, \dots, k$; $j = 1, 2, \dots, n_i$ болгон хөрвүүлнэ. Дундаж утгыг олвол $\bar{z}_i = a'\bar{y}_i$ гарах ба үүнд $\bar{y}_i = \sum_{j=1}^{n_i} y_{ij} / n_i$ байна. Хоёр бүлэгтэй тохиолдолд $\bar{z}_1, \bar{z}_2, \dots, \bar{z}_k$ хамгийн сайн тусгаарлаж чадахаар a векторыг хайдаг.

$a'(\bar{y}_1 - \bar{y}_2) = (\bar{y}_1 - \bar{y}_2)'a$ байх тул бид (2.1) тэгшитгэлийг дараах байдлаар илэрхийлж болно.

$$\frac{(\bar{z}_1 - \bar{z}_2)^2}{s_z^2} = \frac{[a'(\bar{y}_1 - \bar{y}_2)]^2}{a'S_{pl}a} = \frac{a'(\bar{y}_1 - \bar{y}_2)(\bar{y}_1 - \bar{y}_2)'a}{a'S_{pl}a}. \quad (2.7)$$

k бүлэгт (2.7)-г харуулахын тулд $(\bar{y}_1 - \bar{y}_2)(\bar{y}_1 - \bar{y}_2)'$ MANOVA-аас H матрицыг мөн S_{pl} -ээс E -г орлуулвал

$$\lambda = \frac{a'Ha}{a'Ea}, \quad (2.8)$$

Болох ба дараах хэлбэрээр харуулж болно.

$$\lambda = \frac{SSH(z)}{SSE(z)}, \quad (2.9)$$

Үүнд $H = \frac{n_1 n_2}{n_1 + n_2} (\bar{y}_1 - \bar{y}_2)(\bar{y}_1 - \bar{y}_2)'$, $E = \frac{S_{pl}}{n_1 + n_2 - 2}$ гэж илэрхийлэгдэх ба z -ийн хоорондох (between) болон доторх (within) нийлбэрийн квадратууд тус тус $SSH = n \sum_{i=1}^k (\bar{z}_i - \bar{z}_{..})^2$, $SSE = \sum_{ij} (z_{ij} - \bar{z}_{i.})^2$ байна.

(2.8) тэгшитгэлийг хувирган дараах байдлаар илэрхийлж болно.

$$a'Ha = \lambda a'Ea,$$

$$a'(Ha - \lambda Ea) = 0. \quad (2.10)$$

λ болон a -ын утгуудыг шалгахад (2.10) дах шийдүүдээс хамгийн их λ -ын үр дүнг өгдөг a -ын утгуудыг илрүүлж олдог. Хэрэв (2.8)-д өгөгдсөнөөр $\lambda = 0/0$ болоход $a' = 0'$ шийд биш болдог. Бусад шийдүүд нь доорх хэлбэрээр илэрхийлэгдэх ба

$$(Ha - \lambda Ea) = 0, \quad (2.11)$$

Дараах хэлбэрээр мөн илэрхийлж болно.

$$(E^{-1}H - \lambda I)a = 0. \quad (2.12)$$

(2.12) тэгшитгэл дээрх шийдийн хувийн утгууд (eigenvalues) нь $\lambda_1, \lambda_2, \dots, \lambda_s$ ба холбогдох $E^{-1}H$ –ын хувийн векторууд a_1, a_2, \dots, a_s байна. Хувийн утгуудын өмнөх хэлэлцүүлгээс харгалзан үзэн λ -уудыг $\lambda_1 > \lambda_2 > \dots > \lambda_s$ гэж эрэмбэлэсэн. Хувийн утгуудын тоо s нь H -ын эрэмбэ бөгөөд эдгээр утгууд нь p эсвэл $(k-1)$ -ээс бага байна. Хэрэв хамгийн том хувийн утга λ_1 нь (2.8) дахь $\lambda = a'Na/a'Ea$ хамгийн их утга байдаг ба хувийн утга a_1 -д хамаарах хамгийн их утгыг коэффициентын вектор харуулна. Дискриминант функц $z_1 = a'_1u$ дундажийг хамгийн их байхаар тусгаарлана; энд z_1 нь хамгийн их байхаар дундажийг тусгаарлах чиглэл эсвэл хэмжээсийг харуулна.

Дискриминант функц z_i бүрийн харьцангуй ач холбогдол түүний хувийн утгын нийт харьцаагаар тооцно.

$$\frac{\lambda_i}{\sum_{j=1}^s \lambda_j} \quad (2.13)$$

Энэхүү шалгуураар хоёр эсвэл гурван дискриминант функц бүлэг хоорондын ялгааг дүрслэхэд хангалттай байдаг.

2.2.2 Дискриминант функцийн хамаарлын хэмжүүр

Хамааран хувьсагч болох y_1, y_2, \dots, y_p ба бүлгийн үл хамаарах хувьсагч $\mu_i, i = 1, 2, \dots, k$ хоорондын хамаарлыг хэмждэг. Эдгээр хэмжүүрүүд нь хувьсагчид хэр сайн бүлгүүдийг тусгаарлаж чадах вэ гэдэг асуулттай тулгардаг. Рой-ийн статистикийн θ нь R^2 -тай адилтгах хамаарлын хэмжүүр болдог, учир нь энэ нь эхний дискриминант функцийн $z_1 = a'_1u$ болон нийт квадратуудын нийлбэр хоорондох харьцааг харуулдаг.

$$\eta_{\theta}^2 = \theta = \frac{\lambda_1}{1 + \lambda_1} = \frac{SSH(z_1)}{SSE(z_1) + SSH(z_1)}$$

η_{θ}^2 -ийн өөр нэгэн тайлбар нь эхний дискриминант функц болон хамгийн их корреляцийн квадрат хоорондын хамаарал, бүлгийн гишүүдийн хувьсагчид $k-1$ (дамми) хоорондын хамгийн сайн шугаман хослолыг илэрхийлнэ. Хамгийн их корреляцийг (эхний) каноник корреляци гэдэг. Дискриминант функц бүрийн каноник корреляцийн квадрат нь дараах хэлбэрээр бодогдоно.

$$r_i^2 = \frac{\lambda_i}{1 + \lambda_i}, i = 1, 2, \dots, s. \quad (2.14)$$

Энд r_i^2 -ын дундаж $A_p = \frac{\sum_{i=1}^s r_i^2}{s} = \frac{V^{(s)}}{s}$ байна.

2.2.3 Стандартчилагдсан дискриминант функцүүд

Хоёр бүлгийн тохиолдолд тусгаарласан хоёр бүлгийн харьцангуй хувь нэмэр $y's$, дискриминант функцийн коэффициентууд $a_r, r = 1, 2, \dots, p$ -ыг харьцуулан хамгийн сайн үнэлдэг.

$$z = a'y = a_1y_1 + a_2y_2 + a_py_p$$

Хэрэв $y's$ нь тохирч байх үед мэдээллийн чанартай байх ба энэ нь ижил хэмжээст харьцангуй вариациар хэмждэг. Харин $y's$ нь тохирохгүй байвал хувьсагчдаа стандартчилахын тулд a_r^* хувьсагчийг шинээр бий болгодог.

Хоёр бүлгийн тохиолдлыг үзье. Бүлэг нэг эсвэл хоёрт байгаа i -р таамаглалын вектор y_{1i} эсвэл y_{2i} -ын хувьд дискриминант функцийн стандартчилагдсан хувьсагчдыг доорх хэлбэрээр харуулна.

$$\begin{aligned} z_{1i} &= a_1^* \frac{y_{1i1} - \bar{y}_{11}}{s_1} + a_2^* \frac{y_{1i2} - \bar{y}_{12}}{s_2} + \dots + a_p^* \frac{y_{1ip} - \bar{y}_{1p}}{s_p}, \\ z_{2i} &= a_1^* \frac{y_{2i1} - \bar{y}_{21}}{s_1} + a_2^* \frac{y_{2i2} - \bar{y}_{22}}{s_2} + \dots + a_p^* \frac{y_{2ip} - \bar{y}_{2p}}{s_p}, \\ i &= 1, 2, \dots, n_2 \end{aligned} \quad (2.15)$$

Энд хоёр бүлгийн дундаж векторууд нь тус тус $\bar{y}'_1 = (\bar{y}_{11}, \bar{y}_{12}, \dots, \bar{y}_{1p})$, $\bar{y}'_2 = (\bar{y}_{21}, \bar{y}_{22}, \dots, \bar{y}_{2p})$ байх ба r -р хувьсагчийн түүврийн доторх³ стандарт хазайлтыг s_r, s_{pl} -ын r -р диагоналийн элементийн квадрат язгуураар илэрхийлсэн болно. Дараах байдлаар стандартчилагдсан коэффициентүүд харагдана.

$$a_r^* = s_r a_r, r = 1, 2, \dots, p. \quad (2.16)$$

Үүнийг вектор хэлбэрээр бичвэл

$$a^* = (\text{diag} S_{pl})^{1/2} a. \quad (2.17)$$

болох ба олон бүлэгтэй тохиолдолд дискриминант функцийг стандартчилахдаа хоёр бүлгийн тохиолдолтой адил загварыг ашиглана. Хэрэв m -р дискриминант функцийн r -р коэффициентийг a_{mr} , $m = 1, 2, \dots, s; r = 1, 2, \dots, p$ гэж тэмдэглэвэл стандартчилагдсан загвар нь

$$a_{mr}^* = s_r a_{mr},$$

³ доторх- within

болох ба s_r нь $S_{pl} = E/v_E$ диагоналийн бүлэг доторх⁴ стандарт хазайлтыг харуулна. Олон дискриминант функцтэй учраас a_{mr}^* -ыг хоёр индексээр тэмдэглэдэг. Харин хоёр бүлэгт нэг дискриминант функц байдаг учраас a_r^* нэг индекстэй байна.

Өөрөөр хэлбэл m -р хувийн вектор нь скаляраар үржүүлэх хүртэл ганц байх тул стандартчилалыг хялбарчилвал:

$$a_{mr}^* = \sqrt{e_{rr}} a_{mr}, \quad r = 1, 2, \dots, p,$$

хэлбэртэй болох ба e_{rr} нь E -ийн r -р диагональ элемент байна.

2.3 Ач холбогдлын тест

Магадлалыг шалгахдаа олонлогийг хэвийн тархалттай⁵ гэсэн таамаглал тавьдаг. Энэ дискриминант функцийг хөгжилд зайлшгүй байх шаардлагатай биш юм.

2.3.1 Хоёр бүлгийн тохиолдлын⁶ тест

Дискриминант функц $z = a'y$ -ээр олж авсан тусгаар хөрвүүлсэн дундаж $(\bar{z}_1 - \bar{z}_2)^2 / s_z^2$ нь дундаж вектор \bar{y}_1 ба \bar{y}_2 -уудын хоорондох стандартчилагдсан зайтай эквалент байна. Энэхүү стандартчилагдсан зай нь хоёр бүлэг T^2 -тай пропорциональ байна. Дискриминант функцийг коэффициент вектор a нь T^2 ач холбогдолтой үед 0-оос ялгаатай байна. Илүү дэлгэрэнгүй харуулвал, дискриминант функцийг коэффициент вектор нь $\alpha = \Sigma^{-1}(\mu_1 - \mu_2)$, $H_0: \alpha = 0$ нь $H_0: \mu_1 = \mu_2$ -той эквалент байна.

2.3.2 Олон бүлгийн тохиолдлын тест

Дискриминант шалгуур үзүүлэлт $\lambda = a'Na/a'Ea$ нь λ_1 , $E^{-1}N$ -ийн хамгийн том өөрийн утгаар хамгийн дээд цэгтээ хүрдэг бөгөөд үлдсэн өөрийн утгууд $\lambda_2, \dots, \lambda_s$ нь бусад дискриминант хэмжээстэй нийцтэй байна. Энэхүү өөрийн утгууд нь дундаж векторуудын ач холбогдлын ялгааг Вилкийн Λ -тай ижил байна.

$$\Lambda_1 = \prod_{i=1}^s \frac{1}{1 + \lambda_i} \quad (2.18)$$

Үүний тархалт нь $\Lambda_{p,k-1,N-k}$ байх ба үүнд жигд бус загварт⁷ $N = \sum_i n_i$ эсвэл жигд загварт⁸ $N = kn$ байдаг. Хэрэв нэг ба түүнээс дээш λ_i нь том байвал хувийн утгууд ба дискриминант функцийг ач холбогдлыг Вилкийн Λ тестийг ашиглана. s хувийн утгууд нь s хэмжээст тусгаарласан дундаж векторууд $\bar{y}_1, \bar{y}_2, \dots, \bar{y}_k$ -ыг харуулна. Энэхүү хэмжээсүүдийн аль нь ач холбогдолтой байх нь ашигтай. Вилкийн Λ нь MANOVA⁹-ын

⁴ доторх- within

⁵ олонлогийн хэвийн тархалт- multivariate normality

⁶ тохиолдол- case

⁷ жигд бус загвар- unbalanced design

⁸ жигд загвар- balanced case

⁹ MANOVA- ваиацийн шинжилгээ

бусад гурван тестийн статистикаас илүү ач холбогдолтой учир нь үүнийг хувийн утгуудын дэд олонлог дээр ашиглаж болно. Λ -ын критик утгыг тусгай хүснэгтээс харах ба Λ_1 нь χ^2 -тай ойролцоолчлогдон $v_E = N - k = \sum_i n_i - k$ ба $v_H = k - 1$ байж доорх хэлбэрээр бичиж болно:

$$\begin{aligned} V_1 &= - \left[v_E - \frac{1}{2}(p - v_H + 1) \right] \ln \Lambda_1 \\ &= - \left[N - 1 - \frac{1}{2}(p + k) \right] \ln \prod_{i=1}^s \frac{1}{1 + \lambda_i} \\ &= [N - 1 - \frac{1}{2}(p + k)] \sum_{i=1}^s \ln(1 + \lambda_i), \end{aligned} \quad (2.19)$$

Энд чөлөөний зэрэг нь $p(k - 1)$ ба χ^2 -тай ойролцоо байна. Тест статистик Λ_1 мөн (2.19)-ийн ойролцоолчлол нь бүхий л $\lambda_1, \lambda_2, \dots, \lambda_s$ -ын ач холбогдлыг шалгадаг. Хэрэв тест H_0 -ийг няцаавал дор хаяж нэг λ нь тэгээс ялгаатай, түүнчлэн тусгаарласан дундаж векторын нь дор хаяж нэг хэмжээстэй байна. Хэрэв λ_1 хамгийн том бол $z_1 = a'_1 u$ ач холбогдолтой гэж үзэн гэдэгт итгэлтэй байж болно.

$\lambda_2, \lambda_3, \dots, \lambda_s$ ач холбогдлыг шалгаж үзэхийн тулд λ_1 -г Вилкийн Λ -аас хасаж, χ^2 -тай ойролцоолчлон

$$\Lambda_2 = \prod_{i=2}^s \frac{1}{1 + \lambda_i}, \quad (2.20)$$

$$V_2 = - \left[N - 1 - \frac{1}{2}(p + k) \right] \ln \Lambda_2 = [N - 1 - \frac{1}{2}(p + k)] \sum_{i=2}^s \ln(1 + \lambda_i), \quad (2.21)$$

эдгээр нь $(p - 1)(k - 2)$ чөлөөний зэрэгтэй ба χ^2 -тай ойролцоо байна. Хэрэв тест H_0 -ийг няцаавал дор хаяж λ_2 нь дискриминант функц $z_2 = a'_2 u$ -д хамааралтай, ач холбогдолтой байна. Энэхүү загварыг λ_i бүрт H_0 -ийг няцааж чадахгүй болтол шинжилнэ.

m -р алхамын тест статистик нь

$$\Lambda_m = \sum_{i=m}^s \frac{1}{1 + \lambda_i}, \quad (2.22)$$

Үүний тархалт нь $\Lambda_{p-m+1, k-m, N-k-m+1}$. Статистикийн хувьд

$$\begin{aligned} V_m &= - \left[N - 1 - \frac{1}{2}(p + k) \right] \ln \Lambda_m \\ &= [N - 1 - \frac{1}{2}(p + k)] \sum_{i=m}^s \ln(1 + \lambda_i) \end{aligned} \quad (2.23)$$

болох ба чөлөөний зэрэг болох $(p - m + 1)(k - m)$, χ^2 тархалттай ойролцоо байна. Зарим тохиолдолд судлаачид ач холбогдолтой гэж үзсэнээс илүү λ нь статистик ач

холбогдолтой байдаг. Хэрэв $\lambda_i / \sum_j \lambda_j$ жижиг, ач холбогдолтой байвал холбогдох дискриминант функц нь сонирхолгүй (may not be of interest) байж болно.

Λ_i болгонд F -ойролцоолчлолыг¹⁰ ашиглана.

$$\Lambda_1 = \prod_{i=1}^s \frac{1}{1+\lambda_i},$$

Муллер болон Петерсоны үзсэн F ойролцоолчлол ба үүнд $\nu_E = N - k$ ба $\nu_H = k - 1$:

$$F = \frac{1-\Lambda_1^{1/t}}{\Lambda_1^{1/t}} \frac{df_2}{df_1}, \quad (2.24)$$

Үүнд

$$t = \sqrt{\frac{p^2(k-1)^2-4}{p^2+(k-1)^2-5}}, \quad w = N - 1 - \frac{1}{2}(p+k),$$

$$df_1 = p(k-1), \quad df_2 = wt - \frac{1}{2}[p(k-1) - 2].$$

учир нь

$$\Lambda_m = \prod_{i=m}^s \frac{1}{1+\lambda_i}, \quad m = 2, 3, \dots, s,$$

Дараах томъёог ашиглана.

$$F = \frac{1-\Lambda_m^{1/t}}{\Lambda_m^{1/t}} \frac{df_2}{df_1}, \quad (2.25)$$

p ба $(k-1)$ – ын оронд $(p-m+1)$ болон $(k-m)$ – г орлуулан доорх хэлбэртэй болно.

$$t = \sqrt{\frac{(p-m+1)^2(k-m)^2-4}{(p-m+1)^2+(k-m)^2-5}}$$

$$w = N - 1 - \frac{1}{2}(p+k),$$

$$df_1 = (p-m+1)(k-m),$$

$$df_2 = wt - \frac{1}{2}[(p-m+1)(k-m) - 2]$$

2.4 Дискриминант функцийн тайлбар

Хувьсагч болгоны хувь нэмэр болон дискриминант функцийн тайлбар хооронд маш ойрхон зохилцолтой байдаг. Тайлбарлахдаа коэффициентүүдийн тэмдэгийг харгалзан

¹⁰ ойролцоолчлол- dimensional

үзнэ: хувь нэмрийг тодорхойлохдоо, тэмдэглэгээг орхигдуулахгүй коэффициентууд нь үнэмлэхүй үнэлэгдсэн байдаг. Функцийг тайлбарлахаас илүү хувьсагчдын оруулсан хувь нэмрийг тодорхойлоход илүү анхаарлаа хандуулдаг.

Дараагийн гурван хэсэгт бүлгүүдийг тусгаарлахдаа хувьсагч бүрийн (бусад хувьсагчдын хамт) хувь нэмрийг үнэлхэд нийтлэг гурван аргыг хамардаг. Энэхүү гурван арга нь (1) стандартчилагдсан дискриминант функцийг шалгах, (2) хувьсагч бүрт хэсэгчлэн F -тест тооцоолох (3) дискриминант функц болон хувьсагч бүрийн хоорондох корреляцийг тооцох аргууд байна. 3 дахь арга нь хамгийн их ашиглагддаг. Гуравдугаар аргыг хамгийн ихээр санал болгодог хэдий ч хамгийн бага ашиг тустай байдаг байна.

2.4.1 Стандартчилагдсан коэффициент

Хувьсагчдаас ялгаатай хэмжээсийг нөхөж авахын тулд дискриминант функцийг коэффициентуудыг (2.16) эсвэл (2.17)-г ашиглан стандартчилах боломжтой ба энд коэффициентууд нь стандартчилагдсан хувьсагчдыг хамаарахаар тохируулсан байна. Хоёр бүлгийн эхэнд хийгдсэн ажиглалтад (2.15)-д үзүүлсэнчлэн

$$z_{1i} = a_1^* \frac{y_{1i1} - \bar{y}_{11}}{s_1} + a_2^* \frac{y_{1i2} - \bar{y}_{12}}{s_2} + \dots + a_p^* \frac{y_{1ip} - \bar{y}_{1p}}{s_p},$$

$$i = 1, 2, \dots, n_1.$$

Стандартчилагдсан хувьсагчид $(y_{ir} - \bar{y}_{1r})/s_r$ нь чөлөөт хэмжээстэй ба стандартчилагдсан коэффициент $a_r^* = s_r a_r$, $r = 1, 2, \dots, p$ -тай байна, иймээс бүлгүүдийг хамгийн ихээр тусгаарлахын тулд дискриминант функц z -ын хувьсагчдын хувь нэмрийг зөв тусгаж чаддаг. Олон бүлгийн тохиолдолд дискриминант функц бүрийн коэффициентийн утга $a = (a_1, a_2, \dots, a_p)'$ нь $E^{-1}N$ -ын хувийн вектор байх ба энэ нь хувьсагчдын дунд түүврийн корреляци мөн хувьсагч бүрийн нөлөөг бусадтай харьцуулан тусгадаг.

Коэффициентуудын үнэмлэхүй утгууд нь хувьсагчдыг бүлэг тус бүрт ангилхад оруулсан хувь нэмрийг ашиглаж болно. Дискриминант функцийг тайлбарлах эсвэл нэрлэх хэрэгтэй болвол тэмдэгүүдийг анхаарч үзэх хэрэгтэй. Жишээ нь $Z_1 = 0.8y_1 - 0.9y_2 + 0.5y_3$ нь $z_2 = 0.8y_1 + 0.9y_2 + 0.5y_3$ -аас ялгаатай утгатай байна. Учир нь эхний тэгшитгэл ялгаварыг харуулж байгаа бол хоёр дох тэгшитгэл нийлбэрийг харуулж байгаа юм.

Дискриминант функц нь регрессийн тэгшитгэл зэрэг бусад шугаман хослолуудтай адил хязгаарлалтуудтай байна. Жишээ нь (1) хувьсагч нэмэгдсэн эсвэл хасагдсан тохиолдолд хувьсагчийн коэффициент мэдэгдэхүйц өөрчлөгдөж болно. Мөн (2) хувьсагчын тоотой харьцуулахад түүврийн хэмжээ харьцангуй их биш бол коэффициентууд нь түүврээс түүврийн хооронд тогтвортой байж чадахгүй. Хязгаарлалт 1-д энэхүү загварын коэффициентуудын хувьсагч бүрт оруулах хувь нэмрийг үздэг. Хязгаарлалт 2-д олон тооны хувьсагчдын боловсруулалт нь чөлөөт байдаггүй. Хэрэв N/p хэтэрхий бага байх нь нэг түүвэрт өндөр эрэмбэнд байгаа хувьсагч нөгөө түүвэрт ач холбогдол багатай байна гэсэн үг юм.

2.4.2 Хэсгийн F -утгууд¹¹

Ямар ч y_r -хувьсагчидын хувьд бусад хувьсагчдыг засварласны дараа хэсгийн F -тестээр y_r -ын ач холбогдлыг тооцоолох ба үүнээс гадна y_r -ийг бусад хувьсагчдаас хамааруулан тусгаарлана. Хувьсагч бүрт хэсгийн F -тестийг тооцоолсоны дараа хувьсагчид эрэмбэлэгдэх боломжтой болно.

Хоёр бүлгийн тохиолдолд хэсгийн F -тест дараах байдлаар өгөгдөнө.

$$F = (v - p + 1) \frac{T_p^2 - T_{p-1}^2}{v + T_{p-1}^2}, \quad (2.26)$$

Энд T_p^2 нь бүх p хувьсагчдыг агуулсан хоёр түүврийн Хотеллинг T^2 юм. T_{p-1}^2 нь y_r ба $(v = n_1 + n_2 - 2)$ -аас бусад хувьсагчдыг оруулсан T^2 статистик юм. Энэхүү F статистик нь $F_{1,v-p+1}$ гэсэн тархалттай байна.

Харин олон бүлэгтэй тохиолдолд хэсгийн Λ -г $(p - 1)$ -ээс бусад хувьсагчдад y_r -ийг засварласнаар ашиглана.

$$\Lambda(y_r | y_1, \dots, y_{r-1}, y_{r+1}, \dots, y_p) = \frac{\Lambda_p}{\Lambda_{p-1}}, \quad (2.27)$$

Үүнд Λ_p нь y_r -аас бусад бүх p хувьсагчид болон Λ_{p-1} -г оруулсан Вилкийн Λ -ын утга утга юм. Үүнд хэсгийн F нь дараах байдлаар өгөгднө.

$$F = \frac{1 - \Lambda}{\Lambda} \frac{v_E - p + 1}{v_H}, \quad (2.28)$$

(2.27)-д өгөгдсөн Λ -ын хувьд $v_E = N - k$ ба $v_H = k - 1$ байна. Уг тэгшитгэлд өгөгдсөн хэсгийн Λ статистик нь Λ_{1,v_H,v_E-p+1} гэсэн тархалттай байх ба хэсгийн F нь F_{v_H,v_E-p+1} хэмээх тархалттай байна.

(2.26) ба (2.27)-ын хэсгийн F утгууд нь тусгаар бүлгийн нэг хэмжээс хамааралгүй харин стандартчилагдсан дискриминант функцийн коэффициентуудтэй хамааралтай юм. Тухайлбал y_2 нь s дискриминант функц бүрт ялгаатай хувь нэмэр оруулдаг гэвч тусгаарласан бүлгийн бүх хэмжээсийг тооцоолохын тулд хэсгийн F -ээр y_2 дахь хувь нэмрийн ерөнхий индексийг бүрдүүлдэг. Гэсэн хэдий ч хэсгийн F утгууд хувьсагчдыг эхний дискриминант функцийн коэффициентуудтай ижил дарааллаар эрэмбэлэдэг, ялангуяа $\lambda_i / \sum_j \lambda_j$ том байх тусам эхний функц нь боломжийн тусгаарлалтыг тооцдог.

y_r -ийн нийт хэмжүүртэй ижил хэсгийн индекс нь MANOVA $\eta^2 = \eta_{\Lambda}^2 = 1 - \Lambda$ байх ба доорх хэлбэрээр тодорхойлогдож болно.

$$R_r^2 = 1 - \Lambda_r, \quad r = 1, 2, \dots, p, \quad (2.29)$$

¹¹ Хэсгийн F утгууд- Partial F values

Энд Λ_r нь y_r -ийн хэсэгчилсэн Λ юм. Энэхүү хэсгийн R^2 нь y -ийн $(p - 1)$ -г засварласны дараах y_i болон бүлгийн хувьсагчдын хамаарлыг харуулна.

2.5 Хувьсагчид болон дискриминант функцийн хоорондох хамаарал

Маш олон ном, судалгааны ажлуудад хувьсагчдын ач холбогдлын хамгийн сайн хэмжүүр нь дискриминант функц ба хувьсагч бүрийн хоорондын хамаарал $r_{y_j z_j}$ гэсэн байдаг. Эдгээр хамаарлууд нь дискриминант функцийн хувьсагчдын хувь нэмрийг харгалзсан стандартчилагдсан коэффициентуудаас илүү мэдээлэл өгнө гэж үздэг. Корреляцууд нь хувьсагч бүрийн хувьд t эсвэл F -ыг үржүүлдэг бөгөөд бусад хувьсагчдын нөлөөг тооцохгүйгээр тухайн хувьсагч бүр өөрөөрөө бүлгүүдийг хэрхэн тусгаарлахыг үздэг. Тиймээс эдгээр корреляцууд нь хувьсагчдыг тусгаарлахад ямар хувь нэмэр оруулдаг талаарх мэдээлэл байдаггүй.

Үүнийг ойлгосноор олон талт мэдээллээр хангах корреляцуудын доголдлыг урьдчилан таамаглаж болох байсан. Стандартчилагдсан коэффициентын дутагдал нь зарим хувьсагчдыг хасагдах, зарим нь нэмэгдэхэд өөрчлөгддөг учир тэдгээр нь “тогтворгүй” болдог гэсэн таамаглал дээр тулгуурладаг. Гэхдээ хувьсагч хоорондын харилцан нөлөөг тусгахын тулд дээрх аргаар байсан нь зүгээр байдаг. Олон хүчин зүйлийн шинжилгээгээр үзэж буй загварын хувьсагчдын багцын гүйцэтгэл дээр төвлөрдөг. Бусад бүх хувьсагчдаас үл хамаарах хувьсагч бүрийн оруулах хувь нэмрийг судлах нь бусад хувьсагчдыг хасах нэгдмэл индексийг шаардах явдал юм. Корреляци $r_{y_i z_j}$ нь хувьсагчдыг нэмж эсвэл хасаагүй тохиолдолд тогтвортой байдаг.

2.6 Цэгэн Түгэлт¹²

Дискриминант шинжилгээнд нөлөөлж буй хэмжигдэхүүнийг хэмжээсийн хувьд багасгах ач холбогдол нь дүрслэх¹³ боломж юм. $E^{-1}H$ -ын том хувийн утгууд нь дундаж векторуудын байрласан зайны хэмжээсийг тодорхойлдог. Олон өгөгдлийн утгуудад эхний хоёр дискриминант функц нь $\lambda_1 + \lambda_2 + \dots + \lambda_s$ -ыг тооцоолох ба ингэснээр дундаж векторын загварыг хоёр хэмжээст талбайд үр дүнтэй дүрслэн үзүүлэх боломжтой. Хэрэв хэмжээс 2-оос дээш өгөгдвөл хоёр хэмжээст талбайн бүлэг дотор гажуудал үүсэх боломжтой юм. Өөрөөр хэлбэл хоёр хэмжээст зайд давхцаж буй зарим бүлгүүд гурав дахь хэмжээсээр сайн тусгаарлагдах боломжтой.

y_{ij} ажиглалтын векторын эхний хоёр дискриминант функцийн дүрслэл нь $z_{1ij} = a'_{1i} y_{ij}$ ба $z_{2ij} = a'_{2i} y_{ij}$ байна. Үүнд $i = 1, 2, \dots, k; j = 1, 2, \dots, n_i$ гэж тооцоологдох ба цэгэн түгэлтийн утга нь $N = \sum_i n_i$

$$z_{ij} = \begin{pmatrix} z_{1ij} \\ z_{2ij} \end{pmatrix} = \begin{pmatrix} a'_{1i} y_{ij} \\ a'_{2i} y_{ij} \end{pmatrix} = \begin{pmatrix} a'_{1i} \\ a'_{2i} \end{pmatrix} y_{ij} = A y_{ij}. \quad (2.30)$$

¹² Цэгэн түгэлт- Scatter plot

¹³ Дүрслэх- plotting

Хөрвүүлсэн дундаж векторууд нь

$$\bar{z}_i = \begin{pmatrix} \bar{z}_{1i} \\ \bar{z}_{2i} \end{pmatrix} = \begin{pmatrix} a'_1 \\ a'_2 \end{pmatrix} \bar{y}_i = A \bar{y}_i, \quad i = 1, 2, \dots, k \quad (2.31)$$

хувьсагчийн утгын хамт z_{ij} -тэй дүрслэгдэнэ. Зарим тохиолдолд дүрслэл нь зөвхөн хөрвүүлсэн дундаж векторууд $\bar{z}_1, \bar{z}_2, \dots, \bar{z}_k$ -г дүрслэн харуулдаг.

$E^{-1}N$ -ын хувийн утгууд дундаж векторын хэмжээсийг харуулдаг болохоос цэгүүдийг харуулдаггүй хэмээн үздэг. Хувийн ажиглалтын хэмжээс нь p байна. Гэвч хувьсагчууд хоорондоо хамааралтай учраас чухал хэмжээс нь бага байж болдог. Хэрэв $s = 2$ байвал тухайлбал дундаж векторууд зөвхөн хоёр хэмжээст байна. Хувийн таамаглалын векторууд хоёроос илүү хэмжээст зурагддаг. Тэднийг талбайд дүрслэх нь дундаж векторуудыг хоёр хэмжээст хавтгайд төсөөллийг бий болгодог. Тэдний талбайд оруулах нь дундаж векторуудын хоёр хэмжээст хавтгайд төсөөлөл болгодог.

2.7 Хувьсагчдыг сонгох шатлал¹⁴

Олон судалгааны ажлуудад их хэмжээний хамааран хувьсагчууд байдаг бөгөөд туршилт гүйцэтгэгч нь бүлгүүдийг тусгаарлахад зориулж нөөцлөгдсөн тэдгээрийг хасахыг хүсдэг. Энд хувьсагчдын нэг нэгээр нь хасах эсвэл нэмэх замаар хязгаарлагддаг.

Нэмэх сонголт¹⁵ хийхдээ бүлгийг хамгийн их тусгаарлаж чадах нэг хувьсагчаар эхэлдэг. Дараа нь алхам бүрийг Вилкийн Λ -д тулгуурласан хэсгийн F статистикийг хамгийн их болгодог нэг хувьсагч байх ба иймд бусад хувьсагчдын бүлгүүдийн хамгийн их хэмжээгээр дээгүүр болон догуур тусгаарласны тодорхойлно. Алхам бүр дээр хамгийн их хэсэгчилсэн F -тэй хувьсагчийг сонгоно. Энэ үед хамгийн их F_α -аас хэтрэхгүй F -ийн эзлэх хувь α -аас их байна.

Хасах сонголт¹⁶ нь бүх хувьсагчдыг хамруулан дараа нь хэсгийн F -ээс харан хувьсагчдад хамгийн бага нөлөө үзүүлж байгаагаас нь эхлээд хасдаг.

Хувьсагчдыг сонгох шатлал нь хувьсагчдыг нэмэх болон хасах үйлдлийн нийлэмжээс бий болдог. Хувьсагчдыг нэмэх алхам бүрд өмнөх нэмсэн хувьсагчдын нөлөөллийг шалгаж үздэг. Энэхүү шат нь хамгийн том хэсгийн F -д урьдчилан тогтоосон босго утгаас давсан бол зогсоно.

¹⁴ Хувьсагчдыг сонгох шатлал- Stepwise selection of variables

¹⁵ нэмэх сонголт- forward selection

¹⁶ хасах сонголт- backward elimination

III БҮЛЭГ. ЭМПИРИК СУДАЛГАА

Суралцагчдын амжилт нь элсэлтийн үйл явц, шалгуурт нөлөөлөх төдийгүй суралцагчийн сурах хандлагад нөлөөлөх боломжтой тул суралцагчдын амжилтыг урьдчилан таамаглах нь практик талаасаа маш ач холбогдолтой байдаг. Боловсролын байгууллагуудад ялангуяа дээд боловсрол эзэмшиж байгаа эрдэм шинжилгээний гүйцэтгэлийг урьдчилан таамаглах асуудлын талаар олон судалгаа байдаг. Тэдгээрт дийлэнхдээ өмнөх боловсролын дүн, тестийн оноо, ажлын туршлага, нас, хүйс, гэх зэргийг ашиглан шинжилгээ хийсэн байдаг билээ. Үүнтэй адилаар энэхүү судалгааны хүрээнд *Microsoft Excel 2013* програмын өгөгдлийн шинжилгээний хэрэгсэл болох “*Realstat*”-ыг ашиглан “СЭЗДС”-ийн 2014-2015 намрын хичээлийн жилийн ENG111 (Англи хэл I) хичээлийн 4 группын суралцагчдийн тоон мэдээлэлд үндэслэн дискриминант шинжилгээ хийсэн ба бусад шинжилгээтэй харьцуулахын тулд логистик регрессийг ашиглав.

Уламжлалт ёсоор судлаачид ихэнхдээ статистикийн загваруудыг (ж.нь., дискриминант шинжилгээ, олон хүчин зүйлийн регресс, бусад регресс) эсвэл сүлжээний графикийг ашиглан их дээд сургуулийн оюутнуудын амжилтыг урьдчилан таамагласан байдаг. Сүүлийн үед хийгдсэн судалгаануудад (*Мукта Палива, Уша А.Кумар, 2009*) дискриминант шинжилгээгээр гаргасан үр дүнг бусад чанарын өгөгдлийг ашигладаг сүлжээний графикуудтай харьцуулан үзэхэд статистикийн ач холбогдолын хувьд зөрүүгүй байсан юм.

Өгөгдлийг цуглуулахдаа оюутны мэдээллийн сангаас 60 оюутны мэдээллийг цуглуулсан бөгөөд тус өгөгдөл дээр дискриминант шинжилгээ хийсэн. Оюутнуудыг ENG 111 хичээлийг явцын шалгалтын дүнгээр хангалттай (30), хангалтгүй (30) гэж хоёр бүлэгт хуваасан. Ийнхүү ялгасны дараа оюутнуудын явцын шалгалтын дүн хангалттай (1) мөн хангалтгүй (0) байхад нөлөөлж болох хувьсагчдыг бий болгосон.

Үүнд тайлбарлагч хувьсагчдаараа дараах хувьсагчдыг авч үзсэн:

Хүснэгт 1: Хувьсагчдын мэдээлэл

№	Хувьсагчид	Тэмдэглэгээ	Агуулга
1	Урьдчилсан шалгалтын дүн	y_1	Тухайн хичээлийн талаар хэр мэдлэгтэй байгааг тодруулах, түвшин тогтоох шалгалт
2	Гэрийн даалгаварын оноо	y_2	Тус хичээлийн хүрээнд гэрт өгсөн дасгал ажлыг хийж цаг тухайд нь үнэлүүлж байгаа байдлыг харуулна.
3	Бие даалтын оноо	y_3	Тухайн хичээлийн хүрээнд багаараа болон бие даан хийх ажилд гаргаж байгаа хүчин чармайлтын дүгнэх үнэлгээ
4	Илтгэлийн оноо	y_4	Тухайн хичээлийн хүрээнд сэдэв сонгон авч, уран яруу, бусдадаа ойлгогдохоор ярьж буй эсэхийг харгалзсан үнэлгээ

5	Семинарын оноо	y_5	Хичээлийн лекцийн хүрээнд олж авсан мэдлэгээ бататган семинарын хичээлд хэрхэн идэвх чармайлттай оролцдог, эсвэл хэрхэн зохион байгуулах чадварыг үнэлсэн үнэлгээ
6	Ирцын оноо	y_6	Хичээлдээ оролцохын тулд гаргаж буй идэвх чармайлт буюу хичээлээ тасалдаг эсэх

Эх сурвалж: Судлаачийн тооцоолол

Доорх хүснэгтэд тайлбарлагч хувьсагчдын тодорхойлогч статистик үзүүлэлтүүдийг харуулав. Үүнд урьдчилсан шалгалтын дундаж утга 2.814, гэрийн даалгаварын дундаж утга 3.617, бие даалтын дундаж утга 3.546, илтгэлийн дундаж утга 3.873, семинарын дундаж утга 2.739, ирцийн дундаж утга 4.779 тус тус байна.

Хүснэгт 2: Тодорхойлогч статистик үзүүлэлтүүд

	Урьдчилсан шалгалт	Гэрийн даалгавар	Бие даалт	Илтгэл	Семинар	Ирц
Дундаж утга	2.814	3.617	3.546	3.873	2.739	4.779
Стандарт хазайлт	1.174	0.940	1.062	1.450	1.819	0.449
Хэлтийлт	-1.138	-0.437	-1.191	-1.650	-0.441	-2.876
Овойлт	1.057	-0.495	1.297	2.043	-1.251	8.312

Эх сурвалж: Судлаачийн тооцоолол

Өгөгдсөн хоёр түүврийн хувьд дундаж векторууд болон ковариацийн матрицууд нь дараах хэлбэртэй байна. Хүснэгт 3-д нэг болон хоёрдугаар түүврүүд хоорондын дундаж утгуудыг авсан бөгөөд хоорондын зөрүүг харуулав. Үүнээс харахад хоёр бүлгийн хоорондын зөрүүний хэмжээ бага байна.

Хүснэгт 3: Түүврийн дундаж векторууд

	\bar{y}_1	\bar{y}_2	Зөрүү
Урьдчилсан шалгалтын дүн	2.7972	2.8306	-0.0333
Гэрийн даалгаварын оноо	3.1000	4.1333	-1.0333
Бие даалтын дүн	3.0422	4.0500	-1.0078
Илтгэлийн оноо	3.5067	4.2400	-0.7333
Семинарын оноо	2.5639	2.9139	-0.3500
Илтгэлийн оноо	4.6735	4.8853	-0.2118

Эх сурвалж: Судлаачийн тооцоолол

$$S_1 = \begin{pmatrix} 1.162 & -0.10 & -0.146 & -0.002 & -0.24 & -0.02 \\ -0.107 & 0.809 & 0.419 & 0.232 & 0.346 & 0.092 \\ -0.146 & 0.419 & 1.125 & 1.186 & 0.943 & 0.439 \\ -0.002 & 0.232 & 1.186 & 2.945 & 1.980 & 0.805 \\ -0.245 & 0.346 & 0.943 & 1.980 & 3.303 & 0.604 \\ -0.021 & 0.092 & 0.439 & 0.805 & 0.604 & 1.404 \end{pmatrix},$$

$$S_2 = \begin{pmatrix} 1.644 & 0.072 & -0.0801 & -0.2730 & -0.176 & -0.032 \\ 0.072 & 0.436 & -1.8 \cdot 10^{17} & -7.9 \cdot 10^{17} & -3.4 \cdot 10^{17} & 3.8 \cdot 10^{17} \\ -0.80 & -1.82 \cdot 10^{17} & 0.5560 & 0.20601 & 1.35001 & 0.0500 \\ -0.273 & -7.90 \cdot 10^{17} & 0.2060 & 1.05601 & 0.79401 & 0.06701 \\ -0.176 & -3.45 \cdot 10^{17} & 0.3500 & 0.79401 & 3.36801 & 0.09901 \\ -0.032 & 3.840 \cdot 10^{17} & 0.0500 & 0.06701 & 0.09921 & 0.03601 \end{pmatrix}$$

Эдгээрээс $H_0: \Sigma_1 = \Sigma_2$ гэсэн таамаглалыг Хүснэгт 4-д *Box's M* тестээр шалгаж үзэв. Ингэхэд *Box's M* тестийн p -утга 0.001-ээс их байгаа тул ковариацийн матрицууд нэгэн төрлийн байдаг гэсэн урьдач нөхцөлийг хангаж байна.

Хүснэгт 4: Боксын М Тест

<i>M</i>	46.11413
<i>F</i>	1.959766
<i>Чөлөөний зэрэг 1</i>	21
<i>Чөлөөний зэрэг 2</i>	13240.79
<i>p-утга</i>	0.005417

Үүний дараагаар Пүүлэд ковариацийн матриц нь дараах хэлбэртэй гарна.

$$S_{pl} = \begin{pmatrix} 1.4028 & -0.017 & -0.112 & -0.137 & -0.041 & 0.0114 \\ -0.017 & 0.6227 & 0.2937 & 0.2030 & 0.4105 & 0.0859 \\ -0.1126 & 0.2937 & 0.8853 & 0.6957 & 0.7759 & 0.4889 \\ -0.1376 & 0.2030 & 0.6957 & 2.0001 & 1.6644 & 0.8722 \\ -0.0412 & 0.4105 & 0.7759 & 1.6644 & 4.8024 & 0.8440 \\ -0.0114 & 0.0859 & 0.4889 & 0.8722 & 0.8440 & 0.7741 \end{pmatrix}$$

Хоёр түүврийн хүрээнд ач холбогдлыг шалгаж үздэг T^2 тест нь хоёр түүврийн түүвэрлэлтийн нэгж бүр дээр хувьсагчийг хэмждэг. $H_0: \mu_1 = \mu_2$ гэсэн урьдач нөхцөлийг тавьдаг. Дараах хүснэгтээс $T^2 = 33.3552 > T_{6,60}^3 = 14.850$ гарсан учир H_0 таамаглалыг няцаах буюу түүврийн дундажууд хоорондоо ялгаатай гэдгийг илтгэж байна.

Хүснэгт 5: Ач холбогдлын тест

	<i>Утга</i>	<i>F</i>	<i>Чөлөөний зэрэг 1</i>	<i>Чөлөөний зэрэг 2</i>	<i>p-утга</i>
<i>Pillai Trace</i>	0.365116	5.079964	6	53	0.00035
<i>Wilk's Lambda</i>	0.634884	5.079964	6	53	0.00035
<i>Hotelling Trace</i>	0.57509	5.079964	6	53	0.00035
<i>Roy's Lg Root</i>	0.57509				
<i>Hotelling's T2</i>	33.35524				
<i>Критик утга</i>	14.85				

Дискриминант функцийн коэффициентын векторуудыг олж, бүлгийг хамгийн сайн тусгаарлаж байгаа шугаман хослолыг илэрхийлэх боломжтой болдог бөгөөд дараах тэгшитгэлээр илэрхийлэхэд:

$$a'y = -0.1129y_1 - 1.3746y_2 - 0.7777y_3 - 0.1525y_4 + 0.1989y_5 + 0.054y_6$$

хэлбэртэй болох бөгөөд бүлгийг тусгаарлахад оруулж буй хувьсагчид тус бүрийн нөлөөллийг Хотеллинг Т квадрат тестээр 95%-ийн итгэх түвшинд шалгаж үзэхэд бүх хувьсагчид критик утгаасаа давсан учир ач холбогдолтой гарч байна.

Хүснэгт 6: Хотеллинг Т квадрат тест

	<i>y1</i>	<i>y2</i>	<i>y3</i>	<i>y4</i>	<i>y5</i>	<i>y6</i>
T2	33.095	18.961	29.139	33.056	31.410	33.341
Чөлөөний зэрэг 1	5	5	5	5	5	5
Чөлөөний зэрэг 2	51.623	53.523	51.170	51.578	52.702	51.722
F	6.1625	3.5308	5.4260	6.1552	5.8488	6.2083
p-утга	0.00015	0.0078	0.0004	0.0002	0.0002	0.0002
F критик утга	2.39660	2.3894	2.3966	2.3966	2.3929	2.3966

3.1 Дискриминант шинжилгээ болон бусад шинжилгээнүүдээр хийсэн ангиллын үр дүн

Хүснэгт 7: Дискриминант шинжилгээний үр дүн

Бодит гишүүнчлэл	Ажиглалтын утга	Урьдчилан таамагласан утга	
		1	2
Хангалтгүй дүн авсан	30	21	9
Хангалттай дүн авсан	30	7	23
Нийт	60	28	32

Хүснэгт 8: Дискриминант шинжилгээний ангиллын хувь

	Зөв ангилал	Нийт	Зөв ангилал (%)	Буруу ангилал (%)
Хангалттай дүн авсан	21	30	0.7	0.3
Хангалтгүй дүн авсан	23	30	0.76	0.23
Нийт	44	60	0.73	0.26

Дээрх хүснэгтүүдээс харахад бүх үл хамаарах хувьсагчдыг хамарсан ангилалын үр дүнд нийт 30 оюутныг хангалтгүй үнэлгээ авна гэж үзэж байсан боловч бодит байдал дээр 21 оюутан хангалтгүй үнэлгээ авч 70%-ийг зөв таамагласан байна. Харин хангалттай үнэлгээг 30 хүүхэд авна гэж таамаглаж байсан боловч 23 нь хангалттай үнэлгээ авсанаар 76%-ийг зөв таамагласан байна.

Хүснэгт 9: Логистик регрессийн ангиллын үр дүн

Бодит гишүүнчлэл	Ажиглалтын утга	Урьдчилан таамагласан утга	
		1	2
Хангалтгүй дүн авсан	30	21	9
Хангалттай дүн авсан	30	9	21
Нийт	60	30	30

Хүснэгт 10: Логистик регрессийн ангиллын хувь

	Зөв ангилал	Нийт	Зөв ангилал (%)	Буруу ангилал (%)
Хангалттай дүн авсан	21	30	0.70	0.30
Хангалтгүй дүн авсан	21	30	0.70	0.30
Нийт	42	60	0.7	0.3

Хүснэгт 9,10-д логистик регрессийг үнэлэн ангилан матрицыг гаргаж ирсэн билээ. Үүнд нийт 30 оюутнуудаас 21 нь хангалтгүй үнэлгээ авч нийт 70%-ийг зөв ангилсан бол хангалттай үнэлгээ авна гэж таамагласан оюутнууд мөн адил 70%-ийг зөв ангилжээ.

Дээрх хоёр ангиллын үр дүнгээс харахад дискриминант шинжилгээний ангиллын үр дүн нь логистик ангиллын үр дүнтэй төстэй байна. Дискриминант шинжилгээний зөв ангиллын үр дүн дунджаар 73%, логистик регрессийн ангиллын үр дүн 70% байна. Үүнээс харахад логистик регрессийн үр дүнтэй 5%-ийн зөрүүтэй байгаа нь бусад шинжилгээтэй үр дүнгийн хувьд таарч байгааг илтгэж байна.

ДҮГНЭЛТ, САНАЛ

Дискриминант шинжилгээг 1960 онд Сэр Роналд Фишер нээн илрүүлсэн байдаг. Энэхүү шинжилгээ нь чанарын хамааран хувьсагч болон үл хамаарах хувьсагчдын хоорондох уялдаа холбоог тодруулдаг статистик арга хэрэгсэл юм. ДШ нь регрессийн шинжилгээ болон вариацийн шинжилгээтэй төстэй боловч хамааран хувьсагчдын шинж чанаруудаас хамааран өөр байдаг. Тус шинжилгээний хувьд чанарын өгөгдлөөр өгөгдсөн хамааран хувьсагчдыг тодорхой шинж чанарыг харгалзан үзэж хэдэн ч бүлэгт хуваан шинжилгээ хийж болдгоороо давуу талтай. Энэ арга зүйг ашиглан судлаачид шинжилгээ хийхийн тулд хувьсагчдын сонголт, тусгаарлах, тайлбарлах гэх мэт асуудлуудтай тулгардаг ба ангиллыг бий болгохын тулд нэг ба түүнээс олон математик тэгшитгэлүүдийг бий болгож дискриминант функцыг үүсгэх шаардлагатай байдаг. Бүлэг хоорондын ангиллыг бий болгохын тулд ялгах хувьсагч ашиглах ба нэг эсвэл түүнээс олон шугаман хослолыг үүсгэдэг. Хэрэв хоёр хувьсагчид төгс хамааралтай байвал тэр хоёр хувьсагчдыг нэгэн зэрэг оруулан үнэлгээ хийх нь боломжгүй юм. Өөр нэгэн нөхцөл бол олон судалгааны ажлуудад хүн амын ковариацийн матрицыг ашиглан бүлэг тус бүрт тэнцүү байх хэрэгтэй гэж үздэг. Иймээс хамгийн их хэрэглэгддэг арга бол “шугаман” дискриминант функц ба энэ нь энгийнээр шугаман хослолыг харуулах арга байдаг.

Ихэнх эмпирик ажлуудад бүтээгдэхүүний тандалт, шашин, бүс нутгийн байрлал, оюутнуудын гүйцэтгэл гэх мэтийг урьдчилан таамаглах зорилготойгоор дискриминант шинжилгээг ашиглан судалгаа хийсэн байна. Уламжлалт ёсоор судлаачид тус ажлуудад статистикийн загварууд (ж.нь., дискриминант шинжилгээ, олон хүчин зүйлийн регресс, бусад регресс) эсвэл сүлжээний графикийг ашиглан их дээд сургуулийн оюутнуудын амжилтыг урьдчилан таамагласан байдаг. Эдгээр судалгаануудаас сурлагын амжилтанд нөлөөлөх хүчин зүйлсийн талаар дэлгэрэнгүй авч үзэв. Үүнд элсэлтийн үйл явц, хичээлийн лекц семинарын орж байгаа байдал, шалгалт зэрэг нь суралцагчдын сурах хандлаганд нөлөөлөх боломжтой байдаг. Ихэвчлэн дээд боловсролын эрдэм шинжилгээний гүйцэтгэлийг урьдчилан таамаглахаар хийсэн олон судалгаа байдаг бөгөөд хувьсагчдаа өмнөх боловсролын дүн, тестийн оноо, ажлын туршлага, нас, хүйс, гэх зэргийг сонгон авч шинжилгээ хийсэн байдаг билээ. Тиймээс тус судалгаануудтай адилаар энэхүү судалгааны хүрээнд *Microsoft Excel 2013* програмын өгөгдлийн шинжилгээний хэрэгсэл болох “*Realstat*”-ыг ашиглан “СЭЗДС”-ийн 2014-2015 намрын хичээлийн жилийн ENG111 (Англи хэл I) хичээлийн 4 группийн суралцагчдийн тоон мэдээлэлд үндэслэн дискриминант шинжилгээ хийсэн ба бусад шинжилгээтэй харьцуулахын тулд логистик регрессийг ашиглав.

Өгөгдлийг цуглуулахдаа оюутны мэдээллийн сангаас 60 оюутны мэдээллийг цуглуулан дискриминант шинжилгээ хийсэн. Оюутнуудыг ENG 111 (Англи хэл I) хичээлийг явцын шалгалтын дүнгээр хангалттай (30), хангалтгүй (30) гэж хоёр бүлэгт хуваасан. Ийнхүү ялгасны дараа оюутнуудын явцын шалгалтын дүн хангалттай (1) мөн хангалтгүй (0) байхад нөлөөлж болох хувьсагчдыг бий болгосон. Үл хамааран хувьсагчдыг урьдчилсан шалгалтын оноо, бие даалтын оноо, гэрийн даалгаварын оноо, илтгэлийн оноо, семинарын оноо, ирцын оноо гэсэн зургаан өгөгдлийг сонгож авсан. Эмпирик ажлын хүрээнд тайлбарлагч хувьсагчдын тодорхойлогч статистик үзүүлэлтүүдийг харахад

урьдчилсан шалгалтын дундаж утга 2.814, гэрийн даалгаварын дундаж утга 3.617, бие даалтын дундаж утга 3.546, илтгэлийн дундаж утга 3.873, семинарын дундаж утга 2.739, ирцийн дундаж утга 4.779 гэсэн утгууд байв.

Шинжилгээ ёсоор түүврүүдийн вариаци болон ковариацийн матрицууд ижил, түүврийн дунджууд хоорондоо ялгаатай гэсэн урьдач таамаглалуудыг шалгасан. Мөн бүлгүүдийг тусгаарлахад хамгийн сайн болон муу хувьсагчдын хувь нэмрийг тодруулахын тулд статистик шалгууруудаар ач холбогдлыг үзэхэд бүх хувьсагчид бүлгийг тусгаарлахад ач холбогдолтой байв. Иймээс бүх хувьсагчдыг хамруулан ангилал бий болгоход дискриминант шинжилгээний ангилалд ангилалын үр дүнд нийт 30 оюутныг хангалгүй үнэлгээ авна гэж үзэж байсан боловч бодит байдал дээр 21 оюутан хангалтгүй үнэлгээ авч 70%-ийг зөв таамагласан байна. Харин хангалттай үнэлгээг 30 хүүхэд авна гэж таамаглаж байсан боловч 23 нь хангалттай үнэлгээ авсанаар 76%-ийг зөв таамагласан байна. Харин логистик регрессийг үнэлэн ангиллын матрицыг харахад нийт 30 оюутнуудаас 21 нь хангалтгүй үнэлгээ авч нийт 70%-ийг зөв ангилсан бол хангалттай үнэлгээ авна гэж таамагласан оюутнууд мөн адил 70%-ийг зөв ангилжээ. Энэхүү хоёр ангиллын үр дүнгээс харахад дискриминант шинжилгээний ангиллын үр дүн нь логистик ангиллын үр дүнтэй төстэй байна. Дискриминант шинжилгээний зөв ангиллын үр дүн дунджаар 73%, логистик регрессийн ангиллын үр дүн 70% байна. Үүнээс харахад логистик регрессийн үр дүнтэй 5%-ийн зөрүүтэй байгаа нь бусад шинжилгээтэй үр дүнгийн хувьд таарч байгааг илтгэж байв. Иймд урьдчилан таамаглал хийх судалгаануудын хувьд бусад арга хэрэгслүүдийг ашиглахын оронд хялбар арга болох дискриминант шинжилгээний аргыг хэрэглэвэл тухайн судалгааны үр дүн, таамаглалд эерэг нөлөө үзүүлэх боломжтой юм.

НОМ ЗҮЙ

- Abayomi A. Akomolafe, G. A. (2015). *Efficiency of Discriminant Analysis in Identifying Performance of Students at Risk*. Ibadan.
- Blaženka Divjak, D. O. (2009). *Prediction of Academic Performance Using Discriminant Analysis*. Cavtat.
- Cheong Hee Park, H. P. (2005). *A Relationship Between Linear Discriminant Analysis and Generalized Minimum Squared Error Solution*.
- Dijun Luo, C. D. (2011). *Linear Discriminant Analysis: New Formulations and Overfit Analysis*. Texas, USA.
- Dr. D.S. Chaubey, D. H. (2016). *Analysis of the Factors Influencing Academic Performance of Student*. Dehradun.
- Edward W. Thomas, M. J. (2014). *Using Discriminant Analysis to Identify Students at Risk*. Atlanta.
- Elisabeta Jaba, D. V. (2007). *Discriminant Analysis in the Study of Romanian Regional*. Spain.
- Erimafa J.T., I. A. (2009). *Application of discriminant analysis to predict the class*. Ekpoma, Nigeria.
- Fernandez, G. C. (2004). *Discriminant Analysis, A Powerful Classification Technique in Data Mining*. Nevada.
- Friday Zinzendoff Okwonu, A. R. (2012). *A Model Classification Technique for Linear Discriminant Analysis for Two Groups*. Pulau Pinang.
- Gareth M James, T. J. (2001). *Functional Linear Discriminant Analysis for Irregularly Sampled Curves*. Stanford.
- Gheorghita DINCA, M. B. (2015). *Using discriminant analysis for credit decision*. Braşov.
- Issahaku Yakubu, R. D. (2017). *Discriminant Analysis of Demand-Side Roadblocks to Financial Inclusion in Northern Ghana*. Northern Ghana.
- Itziar Irigoien, F. M. (2016). *Weighted Distance Based Discriminant Analysis: The R Package WeDiBaDis*. Barcelona.
- Julianti Kasih, S. S. (2012). *Predicting students' final results through discriminant analysis*. Bandung, Indonesia.
- Kamran Etemad, R. C. (1997). *Discriminant analysis for recognition of human*. Maryland.
- Klecka, W. R. (1980). *Discriminant analysis, Social sciences--Statistical methods*. California, Los Angeles: Sage Publications, Inc.

Line Clemmensen, T. H. (2011). *Sparse Discriminant Analysis*. Stanford, USA.

M. Chalikias, G. K. (2009). *Discriminant Analysis: A Case Study of a War Data Set*. Vari Attica.

ÖZKAN, M. (2016). *Discrimination level of students' ratio, number of students per faculty member and article scores indicators according to place of Turkish universities in international ranking systems*. Gaziantep.

Rencher, A. C. (2002). *Methods of Multivariate Analysis*.

Shin, H. (2008). *An extension of Fisher's discriminant analysis for stochastic processes*. Auburn.

Sougata Banerjee, S. P. (2014). *Predicting Consumer Purchase Intention: A Discriminant Analysis Approach*.

Sugiyama, M. (2007). *Dimensionality Reduction of Multimodal Labeled Data by Local Fisher Discriminant Analysis*. Tokyo, Japan.

T. Ramayah, N. H.-C. (2010). *Discriminant analysis: An illustrated example*. Penang, Malaysia.

Zhihua Qiao, L. Z. (2012). *Effective Linear Discriminant Analysis for High Dimensional, Low Sample Size Data*. Texas.

Шинэцэцэг, Б. (2017). *Зээлдэгчдийн зээлжих чадварыг орчин үеийн статистик аргуудыг ашиглан үнэлэх нь*. Улаанбаатар.

ХАВСРАЛТ**Хавсралт 1: Шинжилгээний хэсэгт ашигласан өгөгдөл**

№	шалгалт	pre-test	гэрийн д	бие даалт	Илтгэл	Семинар	ирц
1	0	1.83	2.50	1.33	3	0.000	5
2	0	3.42	3.50	3.33	3	0.000	5
3	0	4.00	2.50	0.00	0	0.000	3.44
4	0	4.33	1.50	2.00	3	0.000	4.27
5	0	3.75	4.45	1.43	0	0.833	3.435
6	0	2.50	3.40	2.87	5	4.167	4.895
7	0	2.08	3.60	3.00	4.1	1.667	4.895
8	0	4.17	2.85	3.03	5	5.000	5
9	0	4.00	3.60	3.43	5	4.500	4.685
10	0	3.00	3.10	3.20	5	4.750	5
11	0	3.33	2.15	3.47	1.6	1.667	5
12	0	3.42	2.35	3.10	5	5.000	5
13	0	1.92	3.20	3.77	5	3.333	5
14	0	2.17	1.80	2.33	1.5	3.333	4.79
15	0	3.92	1.55	3.33	4.4	2.667	5
16	0	2.67	2.10	1.67	4.4	3.333	4.79
17	0	2.50	2.25	3.20	4.6	1.667	4.685
18	0	2.00	4.55	3.10	5	3.333	5
19	0	2.83	3.30	3.87	5	4.167	4.895
20	0	3.42	3.75	4.33	2.8	1.667	4.795
21	0	3.00	4.85	3.33	3.2	3.333	5
22	0	3.33	3.50	4.00	4	2.500	4.375
23	0	3.00	3.05	4.00	5	0.000	5
24	0	1.92	3.35	3.33	3.6	5.000	5
25	0	3.92	3.00	2.00	0	0.000	3.44
26	0	2.33	3.00	3.80	5.00	5.000	5.00
27	0	1.83	4.50	4.33	3.8	3.333	5
28	0	0.00	2.00	1.33	0	0.000	2.815
29	0	0.00	3.80	4.67	3.5	4.167	5
30	0	3.33	3.95	4.67	4.7	2.500	5
31	1	3.42	4.00	4.33	0	0.000	4.585
32	1	2.92	4.00	1.33	3	0.000	4.375
33	1	0.00	4.00	4.00	5	5.000	4.79
34	1	3.50	3.50	2.67	4	3.333	5
35	1	3.25	3.50	3.67	4	3.333	4.895
36	1	3.00	3.50	4.00	4	3.333	5
37	1	0.00	4.00	3.67	4	0.000	4.895
38	1	3.33	5.00	4.33	5	4.167	5
39	1	3.33	3.50	3.67	4	0.000	4.69
40	1	3.33	3.00	3.67	3	0.000	5

41	1	3.42	4.00	4.33	4	1.667	4.895
42	1	4.50	4.50	4.00	5	0.000	4.895
43	1	2.50	4.00	4.33	5	2.500	5
44	1	2.92	4.00	3.00	5	5.000	4.895
45	1	0.00	4.00	4.67	5	3.333	5
46	1	0.00	5.00	5.00	5	0.000	4.79
47	1	3.00	4.00	4.70	4.6	4.167	5
48	1	2.33	2.30	3.97	5	1.667	4.895
49	1	2.42	5.00	4.67	5	5.000	4.27
50	1	3.08	5.00	4.33	4.5	3.333	5
51	1	4.33	4.50	4.33	4	3.750	5
52	1	4.00	4.25	4.33	4.8	4.167	5
53	1	3.33	4.05	3.33	3	3.917	4.69
54	1	3.33	5.00	4.50	5	4.167	5
55	1	3.33	5.00	4.00	4.3	4.167	5
56	1	4.50	4.75	4.67	4	3.917	5
57	1	2.50	4.75	4.00	5	5.000	5
58	1	4.33	4.50	4.67	5	3.333	5
59	1	2.92	4.00	5.00	4	4.167	5
60	1	2.08	3.40	4.33	4	5.000	5

Эх сурвалж: СЭЗИС-ийн “А” багш дээр “Англи хэл I” судалсан оюутнуудын дүн

Хавсралт 2: Өгөгдөлд тулгуурлан хийсэн хүчин зүйлсийн шинжилгээ

Descriptive statistics

	pre-test	гэрийн д	бие даалт	Илтгэл	Семинар	ирц
Mean	2.81	3.62	3.55	3.87	2.74	4.78
Std dev	1.174446	0.940008	1.062316	1.450291	1.819249	0.449082
Skewness	-1.13757	-0.43677	-1.19108	-1.64986	-0.44071	-2.87595
Kurtosis	1.056674	-0.49537	1.296767	2.043011	-1.25058	8.312497

Factor Matrix (rotated Varimax)

	1	2	Commun	Specific
pre-test	0.023218	0.993825	0.988228	0.011772
гэрийн д	-0.56449	-0.05366	0.321528	0.678472
бие даалт	-0.82675	-0.10752	0.695068	0.304932
Илтгэл	-0.84847	-0.04359	0.721797	0.278203
Семинар	-0.68405	0.047278	0.470157	0.529843
ирц	-0.83078	0.066943	0.694676	0.305324
	2.880709	1.010745	3.891454	2.108546

Reproduced Correlation Matrix

	pre-test	гэрийн д	бие даалт	Илтгэл	Семинар	ирц
pre-test	0.988228	-0.06644	-0.12605	-0.06302	0.031104	0.047241
гэрийн д	-0.06644	0.321528	0.472459	0.48129	0.383601	0.465374
бие даалт	-0.12605	0.472459	0.695068	0.706153	0.560451	0.679646
Илтгэл	-0.06302	0.48129	0.706153	0.721797	0.578331	0.701971
Семинар	0.031104	0.383601	0.560451	0.578331	0.470157	0.571458
ирц	0.047241	0.465374	0.679646	0.701971	0.571458	0.694676

Error Matrix

	pre-test	гэрийн д	бие даалт	Илтгэл	Семинар	ирц
pre-test	0.011772	0.058866	0.044174	-0.01275	-0.04553	-0.03312
гэрийн д	0.058866	0.678472	0.081837	-0.19358	-0.13315	-0.23347
бие даалт	0.044174	0.081837	0.304932	-0.14028	-0.18514	-0.06212
Илтгэл	-0.01275	-0.19358	-0.14028	0.278203	-0.0368	0.016958
Семинар	-0.04553	-0.13315	-0.18514	-0.0368	0.52984 3	-0.12523
ирц	-0.03312	-0.23347	-0.06212	0.016958	-0.12523	0.305324

Factor Scores Matrix - Regression Method

	1	2
pre-test	-0.02317	0.985352
гэрийн д	-0.19483	-0.03549
бие даалт	-0.28444	-0.08068
Илтгэл	-0.29401	-0.01658
Семинар	-0.23963	0.068418
ирц	-0.29133	0.092543

Factor Scores Matrix - Bartlett's Method

	1	2
pre-test	-0.02675	1.006096
гэрийн д	-0.09843	0.001621
бие даалт	-0.32048	0.004144
Илтгэл	-0.36124	0.007538
Семинар	-0.1534	0.005056
ирц	-0.3234	0.01103

Factor Scores Matrix - Anderson-Rubin's Method

78.0157	205.5866
205.5866	7134.178

	1	2
--	---	---

pre-test	-0.02806	1.000557
гэрийн д	-0.09771	0.001622
бие даалт	-0.31812	0.004154
Илтгэл	-0.35859	0.007534
Семинар	-0.15228	0.005044
ирц	-0.32103	0.011003

Хавсралт 3: Өгөгдөлд тулгуурлан хийсэн коварианц болон корреляцийн матрицууд

SSCP Matrices

T

81.38009	-0.49306	-6.02676	-7.61444	-1.81852	0.439319
-0.49306	52.13333	32.65722	23.14167	25.26944	5.775833
-6.02676	32.65722	66.58243	51.43711	42.79463	17.38161
-7.61444	23.14167	51.43711	124.0973	84.29889	27.62607
-1.81852	25.26944	42.79463	84.29889	195.2704	21.50919
0.439319	5.775833	17.38161	27.62607	21.50919	11.8988

H

0.016667	0.516667	0.503889	0.366667	0.175	0.105917
0.516667	16.01667	15.62056	11.36667	5.425	3.283417
0.503889	15.62056	15.23424	11.08556	5.290833	3.202214
0.366667	11.36667	11.08556	8.066667	3.85	2.330167
0.175	5.425	5.290833	3.85	1.8375	1.112125
0.105917	3.283417	3.202214	2.330167	1.112125	0.6731

E

81.36343	-1.00972	-6.53065	-7.98111	-1.99352	0.333403
-1.00972	36.11667	17.03667	11.775	19.84444	2.492417
-6.53065	17.03667	51.34819	40.35156	37.5038	14.1794
-7.98111	11.775	40.35156	116.0307	80.44889	25.2959
-1.99352	19.84444	37.5038	80.44889	193.4329	20.39707
0.333403	2.492417	14.1794	25.2959	20.39707	11.2257

Group Covariance Matrices

0

1.161726	-0.10675	-0.14564	-0.00205	-0.24522	-0.02098
-0.10675	0.808966	0.419425	0.232241	0.345977	0.091578
-0.14564	0.419425	1.215091	1.185916	0.942995	0.438698
-0.00205	0.232241	1.185916	2.945471	1.980307	0.804838
-0.24522	0.345977	0.942995	1.980307	3.30253	0.603878
-0.02098	0.091578	0.438698	0.804838	0.603878	0.350961

1

1.64391	0.071935	-0.07955	-0.27316	0.176477	0.032472
0.071935	0.436437	0.168046	0.173793	0.338314	-0.00563
-0.07955	0.168046	0.555536	0.205517	0.350239	0.050247

-0.27316	0.173793	0.205517	1.055586	0.793793	0.067434
0.176477	0.338314	0.350239	0.793793	3.367569	0.099469
0.032472	-0.00563	0.050247	0.067434	0.099469	0.036133

Pooled covariance matrix

1.402818	-0.01741	-0.1126	-0.13761	-0.03437	0.005748
-0.01741	0.622701	0.293736	0.203017	0.342146	0.042973
-0.1126	0.293736	0.885314	0.695716	0.646617	0.244472
-0.13761	0.203017	0.695716	2.000529	1.38705	0.436136
-0.03437	0.342146	0.646617	1.38705	3.335049	0.351674
0.005748	0.042973	0.244472	0.436136	0.351674	0.193547

Correlation matrix

1	-0.00757	-0.08187	-0.07577	-0.01443	0.014118
-0.00757	1	0.554296	0.28771	0.250449	0.231902
-0.08187	0.554296	1	0.565869	0.375311	0.617531
-0.07577	0.28771	0.565869	1	0.54153	0.718929
-0.01443	0.250449	0.375311	0.54153	1	0.446225
0.014118	0.231902	0.617531	0.718929	0.446225	1

Хавсралт 4: Өгөгдөлд тулгуурлан хийсэн нэг чиглэлт ANOVA-н үр дүн

ANOVA: Single Factor

pre-test

DESCRIPTION					Alpha	0.05		
Groups	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper
0	30	83.916666 67	2.79722 2	1.16172 6	33.6900 5	0.21624 2	2.36436 7	3.23007 8
1	30	84.916666 67	2.83055 6	1.64391	47.6733 8	0.21624 2	2.3977	3.26341 1

ANOVA								
Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0.017	1	0.01666	0.01188	0.91357	4.00687	0.0199	-
Within Groups	81.36	58	1.40281					0.01674
Total	81.38	59	1.37932					

ANOVA:
Single Factorгэрийн
д

DESCRIPTION					Alpha	0.05		
Groups	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper

0	30	93	3.1	0.80896 6	23.46	0.14407 2	2.81160 9	3.38839 1
1	30	124	4.13333 3	0.43643 7	12.6566 7	0.14407 2	3.84494 2	4.42172 5

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	16.02	1	16.01667	25.72127	4.33E-06	4.006873	0.925946	0.291795
Within Groups	36.12	58	0.622701					
Total	52.13	59	0.883616					

ANOVA:
Single Factorбие
даалт

DESCRIPTION

Alpha 0.05

Groups	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper
0	30	91.2666667	3.042222	1.215091	35.23763	0.171786	2.698355	3.386089
1	30	121.5	4.05	0.555536	16.11056	0.171786	3.706133	4.393867

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	15.23	1	15.23424	17.20773	0.000111	4.006873	0.757358	0.212678
Within Groups	51.35	58	0.885314					
Total	66.58	59	1.128516					

ANOVA:
Single Factor

Илтгэл

DESCRIPTION

Alpha 0.05

Groups	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper
0	30	105.2	3.506667	2.945471	85.41867	0.258233	2.989757	4.023576
1	30	127.2	4.24	1.055586	30.612	0.258233	3.72309	4.75691

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	8.067	1	8.066667	4.032267	0.049303	4.006873	0.366618	0.048107
Within Groups	116	58	2.000529					
Total	124.1	59	2.103345					

ANOVA:
Single FactorСемина
р

DESCRIPTION

Alpha

0.05

Groups	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper
0	30	76.91666667	2.563889	3.30253	95.77338	0.333419	1.896478	3.2313
1	30	87.41666667	2.913889	3.367569	97.65949	0.333419	2.246478	3.5813

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	1.837	1	1.8375	0.550966	0.460919	4.006873	0.13552	-0.00754
Within Groups	193.4	58	3.335049					
Total	195.3	59	3.309667					

ANOVA:
Single Factor

урц

DESCRIPTION

Alpha

0.05

Groups	Count	Sum	Mean	Variance	SS	Std Err	Lower	Upper
0	30	140.205	4.6735	0.350961	10.17786	0.080322	4.512719	4.834281
1	30	146.56	4.885333	0.036133	1.047847	0.080322	4.724552	5.046114

ANOVA

Sources	SS	df	MS	F	P value	F crit	RMSSE	Omega Sq
Between Groups	0.673	1	0.6731	3.477717	0.067262	4.006873	0.340476	0.039658
Within Groups	11.23	58	0.193547					

Total 11.9 59 0.20167
5

Хавсралт 5: Өгөгдөлд тулгуурлан хийсэн дискриминант шинжилгээний ангиллын үр дүн

1	2	category	1	2	category
0.946919	0.053081	1	0.214972	0.785028	2
0.440865	0.559135	2	0.648878	0.351122	1
0.981326	0.018674	1	0.450916	0.549084	2
0.967256	0.032744	1	0.713952	0.286048	1
0.592309	0.407691	1	0.538344	0.461656	1
0.739901	0.260099	1	0.483523	0.516477	2
0.562759	0.437241	1	0.275409	0.724591	2
0.845469	0.154531	1	0.083676	0.916324	2
0.55357	0.44643	1	0.33765	0.66235	2
0.785125	0.214875	1	0.551025	0.448975	1
0.895554	0.104446	1	0.186355	0.813645	2
0.918689	0.081311	1	0.070456	0.929544	2
0.62263	0.37737	1	0.212061	0.787939	2
0.982655	0.017345	1	0.565336	0.434664	1
0.944219	0.055781	1	0.251518	0.748482	2
0.974449	0.025551	1	0.027841	0.972159	2
0.860553	0.139447	1	0.232311	0.767689	2
0.297462	0.702538	2	0.756758	0.243242	1
0.591091	0.408909	1	0.080942	0.919058	2
0.277778	0.722222	2	0.076715	0.923285	2
0.215073	0.784927	2	0.146672	0.853328	2
0.408421	0.591579	2	0.19786	0.80214	2
0.403437	0.596563	2	0.47789	0.52211	2
0.775832	0.224168	1	0.074258	0.925742	2
0.84855	0.15145	1	0.11636	0.88364	2
0.750737	0.249263	1	0.087432	0.912568	2
0.175439	0.824561	2	0.183181	0.816819	2
0.982028	0.017972	1	0.093448	0.906552	2
0.404723	0.595277	2	0.209541	0.790459	2
0.174953	0.825047	2	0.578024	0.421976	1

Хавсралт 6: Өгөгдөлд тулгуурлан хийсэн логистик регрессийн үр дүн

	<i>coeff b</i>	<i>s.e.</i>	<i>Wald</i>	<i>p-value</i>	<i>exp(b)</i>	<i>lower</i>	<i>upper</i>
<i>Intercept</i>	-9.75259	6.873661	2.013094	0.155947	5.81E-05		
pre-test	0.156132	0.306831	0.258932	0.610855	1.168981	0.640667	2.132958
гэрийн д бие	1.460456	0.535496	7.438142	0.006386	4.307925	1.508179	12.30505
даалт	0.751106	0.505173	2.210662	0.13706	2.119342	0.787403	5.704332
Илтгэл	0.25741	0.382543	0.452784	0.501016	1.293576	0.611182	2.737873
Семинар	-0.21172	0.186918	1.282995	0.257343	0.80919	0.560978	1.167228

ирц 0.095559 0.773497 0.015262 0.901678 1.100274 0.241601 5.01075

Classification Table

	Suc-Obs	Fail-Obs	
Suc-Pred	21	9	30
Fail-Pred	9	21	30
	30	30	60

Accuracy

0.7	0.7
-----	-----

 0.7

Cutoff

0.5

Хавсралт 7: Өгөгдөлд тулгуурлан ангиллын шинжилгээгээр тооцоолсон үр дүн

a	=	-0.137	-1.396	-0.951	-0.223	0.178	0.620	5.448
0	-0.252	-3.491	-1.268	-0.668	0.000	3.100	2.075	0
0	-0.470	-4.887	-3.170	-0.668	0.000	3.100	5.156	0
0	-0.550	-3.491	0.000	0.000	0.000	2.133	0.808	0
0	-0.596	-2.095	-1.902	-0.668	0.000	2.647	1.421	0
0	-0.515	-6.214	-1.363	0.000	0.148	2.130	4.784	0
0	-0.344	-4.748	-2.727	-1.113	0.740	3.035	4.468	0
0	-0.286	-5.027	-2.853	-0.912	0.296	3.035	5.175	0
0	-0.573	-3.980	-2.885	-1.113	0.888	3.100	3.416	0
0	-0.550	-5.027	-3.265	-1.113	0.799	2.905	5.151	0
0	-0.412	-4.329	-3.044	-1.113	0.844	3.100	4.129	0
0	-0.458	-3.002	-3.297	-0.356	0.296	3.100	2.801	0
0	-0.470	-3.282	-2.948	-1.113	0.888	3.100	2.885	0
0	-0.263	-4.469	-3.583	-1.113	0.592	3.100	5.208	0
0	-0.298	-2.514	-2.219	-0.334	0.592	2.970	1.207	0
0	-0.538	-2.164	-3.170	-0.979	0.474	3.100	2.202	0
0	-0.367	-2.932	-1.585	-0.979	0.592	2.970	1.568	0
0	-0.344	-3.142	-3.044	-1.024	0.296	2.905	3.665	0
0	-0.275	-6.354	-2.948	-1.113	0.592	3.100	6.448	1
0	-0.389	-4.608	-3.678	-1.113	0.740	3.035	5.234	0
0	-0.470	-5.237	-4.121	-0.623	0.296	2.973	6.242	1
0	-0.412	-6.773	-3.170	-0.712	0.592	3.100	6.550	1

0	-0.458	-4.887	-3.804	-0.890	0.444	2.713	5.967	1
0	-0.412	-4.259	-3.804	-1.113	0.000	3.100	5.664	1
0	-0.263	-4.678	-3.170	-0.801	0.888	3.100	4.398	0
0	-0.538	-4.189	-1.902	0.000	0.000	2.133	3.420	0
0	-0.321	-4.189	-3.614	-1.113	0.888	3.100	4.607	0
0	-0.252	-6.284	-4.121	-0.846	0.592	3.100	7.307	1
0	0.000	-2.793	-1.268	0.000	0.000	1.745	2.316	0
0	0.000	-5.306	-4.439	-0.779	0.740	3.100	6.683	1
0	-0.458	-5.516	-4.439	-1.046	0.444	3.100	6.998	1
1	-0.470	-5.586	-4.121	0.000	0.000	2.843	6.395	1
1	-0.401	-5.586	-1.268	-0.668	0.000	2.713	4.408	0
1	0.000	-5.586	-3.804	-1.113	0.888	2.970	6.645	1
1	-0.481	-4.887	-2.536	-0.890	0.592	3.100	4.141	0
1	-0.447	-4.887	-3.487	-0.890	0.592	3.035	5.191	0
1	-0.412	-4.887	-3.804	-0.890	0.592	3.100	5.477	1
1	0.000	-5.586	-3.487	-0.890	0.000	3.035	6.928	1
1	-0.458	-6.982	-4.121	-1.113	0.740	3.100	7.918	1
1	-0.458	-4.887	-3.487	-0.890	0.000	2.908	5.899	1
1	-0.458	-4.189	-3.487	-0.668	0.000	3.100	4.786	0
1	-0.470	-5.586	-4.121	-0.890	0.296	3.035	6.797	1
1	-0.619	-6.284	-3.804	-1.113	0.000	3.035	7.548	1
1	-0.344	-5.586	-4.121	-1.113	0.444	3.100	6.932	1
1	-0.401	-5.586	-2.853	-1.113	0.888	3.035	5.228	0
1	0.000	-5.586	-4.439	-1.113	0.592	3.100	7.445	1
1	0.000	-6.982	-4.756	-1.113	0.000	2.970	9.880	1
1	-0.412	-5.586	-4.470	-1.024	0.740	3.100	6.827	1
1	-0.321	-3.212	-3.773	-1.113	0.296	3.035	4.445	0
1	-0.332	-6.982	-4.439	-1.113	0.888	2.647	8.665	1
1	-0.424	-6.982	-4.121	-1.001	0.592	3.100	7.989	1
1	-0.596	-6.284	-4.121	-0.890	0.666	3.100	6.934	1
1	-0.550	-5.935	-4.121	-1.068	0.740	3.100	6.734	1
1	-0.458	-5.655	-3.170	-0.668	0.696	2.908	5.432	0
1	-0.458	-6.982	-4.280	-1.113	0.740	3.100	8.076	1
1	-0.458	-6.982	-3.804	-0.957	0.740	3.100	7.445	1
1	-0.619	-6.633	-4.439	-0.890	0.696	3.100	7.547	1
1	-0.344	-6.633	-3.804	-1.113	0.888	3.100	7.218	1
1	-0.596	-6.284	-4.439	-1.113	0.592	3.100	7.547	1
1	-0.401	-5.586	-4.756	-0.890	0.740	3.100	6.990	1
1	-0.286	-4.748	-4.121	-0.890	0.888	3.100	5.485	1