



САНХҮҮ ЭДИЙН ЗАСГИЙН ИХ СУРГУУЛЬ
ӨДРИЙН ХӨТӨЛБӨР
ЭКОНОМИКСИЙН ТЭНХИМ



Боржигон
Дамдинсүрэнгийн СОЛОНГО

ШАТАЛСАН РЕГРЕССИЙН ОНОЛ АРГА
ЗҮЙН АСУУДЛУУД



Мэргэжлийн индекс
031101

Эдийн засгийн бакалаврын
зэрэг горилсон
Дипломын ажил

Удирдсан
П.ГАНТӨМӨР /Ph.D/

Улаанбаатар. 2020



**САНХҮҮ ЭДИЙН ЗАСГИЙН ИХ
СУРГУУЛЬ
ЭКОНОМИКСИЙН ТЭНХИМ**



**Боржигон
Дамдинсүрэнгийн СОЛОНГО**

**ШАТАЛСАН РЕГРЕССИЙН ОНОЛ АРГА
ЗҮЙН АСУУДЛУУД**

В Мэргэжлийн индекс
031101

**Эдийн засгийн бакалаврын зэрэг
Горилсон дипломын ажил**

Удирдагч: П.Гантөмөр /Ph.D/

Шүүмжлэгч: Д.Түвшинжаргал//

УДИРТГАЛ

Сэдвийн нэр: Шаталсан регрессийн онол арга зүйн асуудлууд

Товч танилцуулга: Уг бакалаврын зэрэг горилсон судалгааны ажлаар шаталсан регрессийн шинжилгээний онол, арга зүйн асуудлыг авч үзсэн. Судалгааны хүрээнд шаталсан регрессийн арга зүйг судлахын тулд эмпирик шинжилгээ хийж, үр дүнгийн тайлбар болон шалгуур үзүүлэлтүүдийн асуудлыг хөндсөн болно. Шаталсан регрессийн шинжилгээг хийхдээ регрессийн шинжилгээний үндсэн таамаглалуудыг дэвшүүлэх бөгөөд загварын сонголт гэсэн нэмэлт үе шат орж ирдэг. Шаталсан регрессийн шинжилгээг хийх арга зүй, үе шатуудыг судалгааны үндсэн хэсэгт дэлгэрэнгүй харуулав.

Эдийн засгийн бүтээлийн сэтгүүлийн ангиллын нэр: E32, C18

Түлхүүр үг: регрессийн шинжилгээ, шаталсан регресс

АГУУЛГА

ОРШИЛ 5

БҮЛЭГ 1. СУДЛАГДСАН БАЙДАЛ 8

1.1 Сэдвийн судлагдсан байдал..... **Error! Bookmark not defined.**

БҮЛЭГ 2. ОНОЛ, АРГА ЗҮЙ 9

БҮЛЭГ 3. ЭМПИРИК ШИНЖИЛГЭЭНИЙ ҮР ДҮН..... 13

3.1 Тодорхойлох статистикийн шинжилгээ..... 13

ДҮГНЭЛТ 29

НОМ ЗҮЙ 30

ХАВСРАЛТ 32

ХҮСНЭГТЭН МЭДЭЭЛЛИЙН ЖАГСААЛТ

Хүснэгт 1. Судлагдсан байдлын тойм.....	6
Хүснэгт 2. Шаталсан регресс ба энгийн регрессийн хоорондын ялгаа	9
Хүснэгт 3. Тодорхойлох статистикийн үзүүлэлтүүд.....	13
Хүснэгт 4. Найдвартай байдлын шинжилгээ, ерөнхий Кронбах альфа коэффициент	14
Хүснэгт 5. Найдвартай байдлын шинжилгээ, дэлгэрэнгүй задаргаа	14
Хүснэгт 6. Корреляцийн шинжилгээний үр дүн.....	17
Хүснэгт 7. Регрессийн үнэлгээний хураангуй	21
Хүснэгт 8. Регрессийн ANOVA тест.....	22
Хүснэгт 9. Регрессийн үнэлгээний коэффициентүүд.....	23
Хүснэгт 10. Үнэлгээнээс хассан хувьсагчид.....	24
Хүснэгт 11. Загварын үлдэгдлийн статистик.....	26

ЗУРГАН МЭДЭЭЛЛИЙН ЖАГСААЛТ

Зураг 1. Регрессийн үлдэгдлийн гистограм	27
Зураг 2. Регрессийн үлдэгдлийн тархалт	28
Зураг 3. Тодорхойлох статистикийн шинжилгээ хийх аргачлал. Алхам 1	32
Зураг 4. Тодорхойлох статистикийн шинжилгээ хийх аргачлал. Алхам 2	32
Зураг 5. Тодорхойлох статистикийн шинжилгээ хийх аргачлал. Алхам 3	32
Зураг 6. Найдвартай байдлын шинжилгээ хийх үе шатууд. Алхам 1	33
Зураг 7. Найдвартай байдлын шинжилгээ хийх үе шатууд. Алхам 2	34
Зураг 8. Найдвартай байдлын шинжилгээ хийх үе шатууд. Алхам 3	34
Зураг 9. Корреляцийн шинжилгээ хийх үе шатууд. Алхам 1.....	35
Зураг 10. Корреляцийн шинжилгээ хийх үе шатууд. Алхам 2.....	36

Зураг 11. Регрессийн шинжилгээ хийх үе шатууд. Алхам 1	36
Зураг 12. Регрессийн шинжилгээ хийх үе шатууд. Алхам 2	37
Зураг 13. Регрессийн шинжилгээ хийх үе шатууд. Алхам 3	38
Зураг 14. Регрессийн шинжилгээ хийх үе шатууд. Алхам 4	39
Зураг 15. Регрессийн шинжилгээ хийх үе шатууд. Алхам 5	40

ХАВСРАЛТЫН ЖАГСААЛТ

Хавсралт 1. Тодорхойлох статистикийн шинжилгээ хийх аргачлал.	32
Хавсралт 2. Найдвартай байдлын шинжилгээ хийх үе шатууд.	33
Хавсралт 3. Корреляцийн шинжилгээ хийх үе шатууд.	35
Хавсралт 4. Регрессийн шинжилгээ хийх үе шатууд	36

ОРШИЛ

Регрессийн шинжилгээ нь аливаа статистикийн шинжилгээнд хувьсагчдын хоорондын хамаарлыг тодорхойлоход ашиглагддаг түгээмэл хэрэгсэл юм. Регрессийн шинжилгээ нь дотроо олон төрөл байдаг, үүнд: логистик регресс, бинар регресс, олон хүчин зүйлийн регресс гэх мэт. Нийгмийн судалгаа, бизнесийн судалгаанд ихээхэн ашиглагддаг төрөл нь олон хүчин зүйлийн регресс бөгөөд судлаач нар ямар хүчин зүйлүүдийг шинжилгээнд оруулах эсэхийг тодорхойлох шаардлагатай байдаг. Одоогоос 50 жилийн өмнө шаталсан регрессийг тайлбарлагч хувьсагчийг сонгох хамгийн шилдэг арга гэж үздэг байсан. Хэдийгээр уг аргачлал нь маш олон шүүмжийг дагуулдаг байсан ч одоог хүртэл хэрэглэгдсээр ирсэн.

Аливаа судалгааны загварт тайлбарлагч хувьсагчийг сонгоход хоёр харилцан эсрэг зорилтыг биелүүлэх хэрэгтэй болдог. Нэгдүгээрт, тухайн загвар нь бодит нөхцөл байдлыг аль болох хамгийн сайнаар тайлбарладаг байх ёстой. Өөрөөр хэлбэл хамгийн бага нөлөөтэй тайлбарлагч хувьсагчийг хүртэл загварт тусгах хэрэгтэй. Хоёрдугаарт, судалгааны загварт аль болох бага хувьсагч ашиглагдах ёстой. Учир нь ач холбогдолгүй хувьсагчид нь регрессийн үнэлгээний үр дүнгийн найдвартай байдлыг болон загварын тайлбарлах чадварыг бууруулна. Аливаа хувьсагчийг сонгох асуудал нь загварыг аль болох хялбар (хувьсагчийн тоог боломжит хамгийн цөөн тоонд хүргэх) байлгах эсвэл бүрэн бүтэн (аль болох олон хувьсагчийг тусгах) байлгах хоёрын хоорондын балансыг олох асуудал болж хувирна. Иймд шаталсан регрессийн аргачлалын онолын болон практик хэрэглээний арга зүйн асуудлыг судлах шаардлага хэрэгцээ байгаагийн үндсэн дээр уг сэдвийг сонгосон болно.

Дипломын ажлын зорилго

Энэхүү судалгааны ажлаар шаталсан регрессийн аргачлалын оновчтой хэрэглээний цар хүрээ, ашиглагдах боломжтой салбар болон арга зүйг нь судлах зорилго тавьсан.

Дипломын ажлын зорилт

Дээрх зорилгын хүрээнд дараах зорилтуудыг дэвшүүллээ:

- Шаталсан регрессийн үүсэл, хөгжил болон түгээмэл хэрэглээний салбаруудыг судлах
- Шаталсан регрессийн давуу болон сул талууд, түүний онолын тавилыг судлах
- Шаталсан регрессийн хандлагууд, үнэлгээний арга, аргачлалын тухай дэлгэрэнгүй судална
- Шаталсан регрессийг орлох боломжтой арга, аргачлалуудыг санал зөвлөмжөөр дэвшүүлэх
-

Судалгааны арга зүй

Уг судалгаанд баримт бичгийн шинжилгээний, индукц-дедукцийн аргачлалыг ашиглах ба эмпирик шинжилгээний хэсэгт тоон шинжилгээний аргачлалыг ашигласан.

Судалгааны ажлын ач холбогдол

Шаталсан регрессийн арга зүйн асуудлыг судалснаар цаашдын эрдэмтэн, судлаач болон оюутнуудад судалгаа шинжилгээний ажил хийх найдвартай хоёрдогч эх үүсвэр болох ба судалгааны арга зүйн суурь болгон ашиглах боломжтой эрдэм шинжилгээний бүтээл болно.

Судлагдсан байдлын тойм

Хүснэгт 1. Судлагдсан байдлын тойм

№	Судлаачдын нэр	Он	Судалгааны ажил	Судалгааны үр дүн
1	Gary Smith	2018	Step away from stepwise	Шаталсан регрессийн гол асуудал бол хамаарах хувьсагчийг тайлбарлахгүй үл хамаарах хувьсагчид нь статистикийн ач холбогдолтой байх магадлалтай. Тус загвар нь түүврийн хувьд тайлбарлах чадамж сайтай байх магадлалтай ч түүврээс гадна буюу эх олонлогийг тайлбарлах чадваргүй байх магадлалтай.
2	T. Johnsson	1992	A procedure for stepwise regression analysis	Уг судалгааны ажлаар туршилтын бус онолын аргачлалаар шаталсан регрессийн арга зүйн асуудлыг судалсан. Уг судалгаа нь шинэ хандлага дээр суурилан ач холбогдол бүхий регрессоруудыг тодорхойлсон.
3	Intan Martina Md Ghani	2010	Stepwise Multiple Regression Method to Forecast Fish Landing	Уг эмпирик судалгаа нь шаталсан регрессийг экологийн салбарт ашиглах туршилтын судалгаа юм.

4	Derksen, Shelley, Keselman, H. J.	1992	Backward, forward and stepwise automated subset selection algorithms: Frequency of obtaining authentic and noise variables	Шаталсан регрессийн хандлагыг ашиглахад түүврийн хэмжээг нэмэгдүүлэх нь загварын найдвартай байдлыг нэмэгдүүлдэггүй.
5	Mitzi Lewis	2007	Stepwise versus Hierarchical Regression: Pros and Cons	Шаталсан регрессийн үр дүнг ашиглахдаа болгоомжтой байх шаардлагатай ба зөвхөн туслах статистик хэрэгсэл болгож ашиглах нь илүү тохиромжтой.

БҮЛЭГ 1. СУДЛАГДСАН БАЙДАЛ

Энэхүү хэсэгт олон улсын судлаачдын хийсэн эмпирик ажлуудыг тоймлон авч үзэх болно.

Гэри Смит 2018 оны (Step away from stepwise) судалгааны ажилдаа загварын найдвартай байдлыг RMSE хэрэгслээр тодорхойлсон ба шаталсан регрессийн гол асуудал бол хамаарах хувьсагчийг тайлбарлахгүй үл хамаарах хувьсагчид нь статистикийн ач холбогдолтой байх магадлалтай гэжээ. Мөн тус загвар нь түүврийн хувьд тайлбарлах чадамж сайтай байх магадлалтай ч түүврээс гадна буюу эх олонлогийг тайлбарлах чадваргүй байх магадлалтай хэмээжээ.

Т. Жонсон 1992 оны (A procedure for stepwise regression analysis) судалгааны ажлаараа туршилтын бус онолын аргачлалаар шаталсан регрессийн арга зүйн асуудлыг судалсан байсан ба уг судалгаа нь шинэ хандлага дээр суурилан ач холбогдол бүхий регрессоруудыг тодорхойлсон юм.

Интан Мартина, Мд Гани нарын 2010 оны (Stepwise Multiple Regression Method to Forecast Fish Landing) судалгааны ажилдаа Малайз улсын 1968-2007 онуудын загасны аж ахуйн үйлдвэрлэлийн өгөгдлийг ашиглан олон хүчин зүйлийн регрессийн шинжилгээг хийсэн бөгөөд ингэхдээ 3 хүчин зүйлээр шаталсан аргачлалаар тооцоолол хийсэн байна. Уг эмпирик судалгаа нь шаталсан регрессийг экологийн салбарт ашиглах туршилтын судалгаа байсан юм.

Дерксен, Шелли, Кеселман, Н. Ж. 1992 оны (Backward, forward and stepwise automated subset selection algorithms: Frequency of obtaining authentic and noise variables) судалгааны ажилд хугацааны цувааны симуляци болох Монте-Карло аргачлалаар загварчлагдсан чимээний хувьсагчыг ашиглан форвард, байквард аргачлалуудыг харьцуулсан байдлаар шинжилсэн бөгөөд шаталсан регрессийн хандлагыг ашиглахад түүврийн хэмжээг нэмэгдүүлэх нь загварын найдвартай байдлыг нэмэгдүүлдэггүй гэсэн дүгнэлтэд хүржээ.

Митзи Льюис 2007 оны (Stepwise versus Hierarchical Regression: Pros and Cons) судалгааны ажилдаа сурагчдын хүн ам зүйн үзүүлэлтүүдийг ашиглан сурагчдын сурлагад нөлөөлөх хүчин зүйлүүдийг тооцсон бөгөөд шаталсан регрессийн үр дүнг ашиглахдаа болгоомжтой байх шаардлагатай ба зөвхөн туслах статистик хэрэгсэл болгож ашиглах нь илүү тохиромжтой хэмээн дүгнэжээ.

БҮЛЭГ 2. ОНОЛ, АРГА ЗҮЙ

Энэ хэсэгт шаталсан регрессийн тухай ойлголт, онолын үндэслэл, үндсэн санаа, үнэлгээний арга аргачлалын талаар авч үзнэ.

Шаталсан регресс нь тайлбарлагч хувьсагчдыг автоматчилагдсан алгоритмын дагуу сонгож, регрессийн загварыг оновчлох статистикийн аргачлал бөгөөд хамгийн бага квадратын аргаар шинжилгээ хийдэг аргачлалуудын нэг юм. Шинжилгээний эхний шатанд хамаарах хувьсагчтай хамгийн их корреляци хамааралтай хувьсагчийг шинжилнэ. Хоёрдугаар шатанд хамгийн их корреляци хамааралтай хувьсагчийг хяналтын хувьсагч болгож, корреляцийн коэффициентээр хоёрдугаарт эрэмбэлэгдэх хувьсагчийг шинжилгээнд оруулна. Уг процессыг бүх хувьсагчуудын хувьд R^2 максимумчлагдах хүртэл ижилхэн хийнэ. Эсвэл уг процессыг эсрэг чиглэлд хийж болно. Үүнд бүх хувьсагчдыг шинжилгээнд оруулж, R^2 -ийн ач холбогдол бүхий өөрчлөлт гарах хүртэл хувьсагчдыг нэг нэгээр нь хасах байдлаар хийнэ.

Шаталсан регрессийн шинжилгээнд ихэвчлэн Фишерийн итгэлцүүрийг ашигладаг:

$$F = \frac{S_1 - S_2}{S_2} \frac{m - n_2}{n_1 - n_2},$$

Үүнд 1 гэсэн индекс нь нэгдүгээр шугаман загварыг илтгэх бол 2 гэсэн индекс нь хоёрдугаар шугаман загварыг илтгэнэ. Хэрэв итгэлцүүрийн утга нь критик утгаас их бол хоёрдугаар загварыг илүү оновчтой гэж үзнэ.

Хүснэгт 2. Шаталсан регресс ба энгийн регрессийн хоорондын ялгаа

	Шаталсан регресс	Олон хүчин зүйлийн регресс
Олон хүчин зүйлийг дэмжих эсэх	Тийм	Тийм
Хэдэн удаа тооцоолол хийдэг	Олон удаа	Нэг удаа
Загварын тайлбарлах чадварыг сонгох үндэслэл	F статистик, Детерминацийн коэффициент	F статистик, Детерминацийн коэффициент
Хувьсагчдын ач холбогдлын түвшин	Хамгийн сайн загварын бүх хувьсагчид ач холбогдолтой байна	Хувьсагчид ач холбогдолгүй эсвэл ач холбогдолтой байх боломжтой
Хувьсагчдын тоо	Хувьсагч нэмж, хасаж м ахмхэм болдог	Хувьсагч нэмэгдэж, хасагдахгүй

2.1.1 Шаталсан регрессийн үндсэн хандлагууд

(Chatterjee et al., 2000) тодорхойлсноор шаталсан регрессийн дараах 3 хандлага байдаг:

Хувьсагч нэмэх хандлага

А шинж чанар бүхий индексийн бүрэлдэхүүнийг авч үзье. Анхдагч бүрэлдэхүүний хувьд аливаа шинж чанарыг авч үзэхгүйгээр зөвхөн тогтмол буюу constant утгыг авч үзнэ. А бүрэлдэхүүн шат бүрд F статистикийг хамгийн их байлгах нөхцөлийг сонгоно:

$$j^* = \arg \max_{j \in J} F_{\text{Add}} \propto \arg \max_{j \in J} \frac{S(A) - S(A \cup \{j\})}{S(A \cup \{j\})}.$$

Хувьсагч хасах хандлага

Хувьсагч хасах хандлагаар анхдагч бүрэлдэхүүн нь бүх шинж чанараас хамаарна гэж таамаглах ба шат бүрт F статистикийн хамгийн бага утгатай шинж чанарыг хасах байдлаар үнэлгээ хийнэ.

$$j^* = \arg \min_{j \in J} F_{\text{Del}} \propto \arg \min_{j \in J} \frac{S(A \setminus \mathbf{x}^j) - S(A)}{S(A)}.$$

Шаталсан регрессийн алгоритм нь Маллоузын тодорхойлсон итгэлцүүр минимумчлагдах нөхцөлд зогсоно:

$$C_p = \frac{S}{MSE} + 2k - m,$$

Үүнд: $MSE = S/n$ – дундаж алдааны квадрат.

Үүнээс гадна Акайке эсвэл Байесын шинжүүрүүдийг ашиглах боломжтой.

$$BIC = \frac{N}{\hat{\sigma}_\epsilon^2} (\overline{err} + (\log N) \frac{d}{N} \hat{\sigma}_\epsilon^2)$$

$$AIC(\alpha) = \overline{err} + 2 \frac{d(\alpha)}{N} \hat{\sigma}_\epsilon^2$$

Хосолсон хандлага.

Уг алгоритмын дагуу хувьсагч бүрийг тухайн шатанд хасах эсвэл нэмэх шийдвэр гаргалтыг хийнэ.

2.1.2 Загварын сонголтын шинжүүр

Шаталсан регрессийн хамгийн өргөн хэрэглэгддэг алгоритмыг хамгийн анх Efronson (1960) санал болгож байсан. Түүний санал болгосон автоматчилалдагдсан алгоритм нь хэтэрхий их тайлбарлагч хувьсагчтай, тэдгээрийн холбоог тайлбарлах суурь онол байхгүй үед ашиглагдах боломжтой байсан. Уг процедурыг ихэвчлэн регрессийн шинжилгээнд ашиглан ба маш олон төрлийн загвар сонголтонд ашиглах боломжтой байж. Түүний аргачлал нь хувьсагч нэмэх хандлагын нэг төрөл байсан. Хувьсагч нэмэх бүрд загварын нийт үлдэгдлийн квадратын нийлбэр дээр суурилан загварын сонголтыг хийх ёстой байж.

2.1.3 Шаталсан загварын онцлог

Шаталсан загварыг ихэвчлэн судалгааны тандах үе шатанд эсвэл прогноз хийх үе шатанд ашиглагдах ба онолыг шалгахын тулд ашиглагддаггүй байна. Онолыг шалгах үе шатанд судлаач хувьсагчдын сонголтыг үндэслэлтэйгээр нотлох бөгөөд компьютерийн алгоритмыг ашиглах нь учир дутагдалтай. Менард (1995, р. 54) бичсэнээр “судлаачдын нийгэмлэгийн нийтээр хүлээн зөвшөөрсөн тохиролцооны дагуу хувьсагч сонгоход компьютерийн шаталсан алгоритм тохиромжгүй. Уг алгоритм нь өгөгдөл дахь санамсаргүй вариацийн хуримтлагдсан дүнг нэмэгдүүлж, ихэвчлэн зөвхөн тухайн түүврийн хувьд загварыг хэт оновчлох бөгөөд уг түүврээс гадна ашиглах боломжгүй үр дүнг бий болгодог”.

Мөн 0.05 ач холбогдлын түвшинд нотлогддог загварын түвшин нь шаталсан регрессийн алгоритмын явцад хамгийн эцсийн шатанд 0.5 түвшинд нотлогдох магадлалтай байдаг ба I төрлийн алдаа үүсгэх магадлалыг нэмэгдүүлдэг (Draper, N.R., Guttman, I. & Lapczak, L., 1979) Иймд Фокс (1991, р. 18) томьёолсны дагуу аливаа загварыг cross-validation ашиглан шалгах хэрэгтэй.

2.1.4 Детерминацийн коэффициентийн өөрчлөлт

Шаталсан регрессийн дагуу аливаа хоёр загварыг хооронд нь харьцуулахад F тест ашигладаг:

$$F = [(R_2^2 - R_1^2)/(k_2 - k_1)]/[(1-R_2^2)/(n - k_2 - 1)],$$

Үүнд: R – загварын детерминацийн коэффициент

K – загвар дахь тайлбарлах хувьсагчдын тоо

n- түүврийн хэмжээ

Фишерийн тестийн дагуу чөлөөний зэргийг ашиглан хоёр загварын детерминацийн коэффициентийн өөрчлөлт нь 0-с ялгаатай эсэхийг шалгана.

2.1.5 Загварын шүүмж

Шаталсан регрессийг ихэвчлэн их хэмжээний өгөгдөлд тандах судалгааны хэрэгслийн хувьд ашиглагддаг.

Wilkinson and Dallal (1981) нарын тодорхойлсноор хувьсагч нэмэх хандлагаар сонгосон загварууд нь 0.1 хувийн түвшинд ач холбогдолтой гэж гарсан боловч бодит байдал дээр Фишерийн статистикийн дагуу дөнгөж 5 хувийн түвшинд ач холбогдолтой байсан байна.

Мөн загварын хэд хэдэн сул тал байдгаас чухал хувьсагч загварт тусгагдахгүй байх магадлалтай ба хоёрдогч шинж чанаруудыг илүүд үзэх явдал юм. Мөн шаталсан регрессийг түүврийн шинж чанартай хэт уялдсанаар детерминацийн коэффициент хэт өндөр гарч, I төрлийн алдаа гарах магадлалыг нэмэгдүүлдэг.

БҮЛЭГ 3. ЭМПИРИК ШИНЖИЛГЭЭНИЙ ҮР ДҮН

Судалгааны энэ хэсэгт банкны ажилтнуудын сэтгэл ханамжид нөлөөлөх хувьсагчдыг шинжлэх ба олон хүчин зүйлийн регрессийн аргачлалыг ашиглана. Мөн шаталсан регрессийн шинжилгээг хийх үе шатууд, арга аргачлалыг дэлгэрэнгүй байдлаар хийнэ.

3.1 Тодорхойлох статистикийн шинжилгээ

Тодорхойлогч статистик нь тухайн түүврийн ерөнхий байр байдал, хувьсагчдын дундаж утгууд, хариултын түвшин болон оролцогчдын үзэл бодлын стандарт хазайлт зэрэг зүйлүүдийг шинжлэх зорилготой тест юм.

Хүснэгт 3. Тодорхойлох статистикийн үзүүлэлтүүд

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
Би ажилдаа дуртай	437	5	9	6.57	.887
Манай байгууллага ажилчдадаа санаа тавьдаг	444	2	10	5.99	2.106
Байгууллагын зүгээс намайг үргэлж дэмжиж байдаг	438	1	10	4.92	2.102
Миний хийдэг ажил сонирхолтой	434	1	9	4.90	1.680
Би хамт ажиллагсдадаа дуртай	447	4	10	7.74	1.632
Миний ажлын байр надад таалагддаг	430	2	10	5.86	1.944
Би цалиндаа сэтгэл ханамжтай байдаг	441	2	10	5.88	2.014
Миний ажлын байрны нөхцөл таатай	443	2	10	5.92	1.945
Миний хийдэг ажил утга учиртай	443	1	10	4.88	1.759
Би хамт ажиллагсадтайгаа таатай харилцаатай	440	3	10	6.97	1.799
Valid N (listwise)	297				

Нийт ажиглалтын 464 утгаас зөвхөн 297 оролцогчийн өгөгдөл хүчинтэй байсан ба эдгээр хүмүүсийн өгөгдөл нь 1 ба түүнээс дээш хариулаагүй үлдээсэн асуулттай байсан.

3.1.1 Найдвартай байдлын шинжилгээ

Найдвартай байдлыг шалгах арга нь тухайн нэг хандлага бүхий тест, хэмжигдэхүүний сонгогдсон түүвэр дэх дотоод нийлэмж, зохицолтой эсэхийг шалгадаг. Хүчин зүйлсийн хоорондох корреляцийн хамаарал их байх тусам Кронбахын альфа өндөр гардаг. Корреляцийн хамаарал их байна гэдэг нь 1 асуултын оноо өндөр байхад нөгөө асуултынх мөн өндөр байх, бага байхад бага байх гэсэн үг юм. Альфа 0-1 хооронд байх бөгөөд хэрэв α 0.6-с илүү байвал тухайн тестийн нийлэмж сайтай гэдэгт итгэж болох бөгөөд α нь өндөр байх тусмаа илүү сайн гэж үздэг. (Moloanto, 2015)

Хүснэгт 4. Найдвартай байдлын шинжилгээ, ерөнхий Кронбах альфа коэффициент

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.724	.748	10

Асуулгын найдвартай байдлын шинжилгээг хоёр байдлаар хийж болдог: ерөнхий Кронбах альфа коэффициент болон хувьсагч тус бүрийн Кронбах альфа коэффициент. Аливаа асуулгын загварыг шалгах бол хувьсагч тус бүрээр задалж харуулсан нь илүү оновчтой байдаг ба шаардлага хангаагүй хувьсагчдыг хасаж, шинжилгээний үр дүнг илүү найдвартай болгоход тустай.

Хүснэгт 5. Найдвартай байдлын шинжилгээ, дэлгэрэнгүй задаргаа

Item-Total Statistics					
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Би ажилдаа дуртай	53.04	83.911	.655	.532	.689
Манай байгууллага ажилчдадаасанаа тавьдаг	53.68	79.285	.306	.346	.717
Байгууллагын зүгээс намайг үргэлж дэмжиж байдаг	54.66	78.482	.346	.399	.709
Миний хийдэг ажил сонирхолтой	54.58	83.232	.331	.172	.710

Би хамт ажиллагсдадаа дуртай	51.90	82.899	.333	.256	.709
Миний ажлын байр надад таалагддаг	53.74	75.368	.478	.398	.685
Би цалиндаа сэтгэл ханамжтай байдаг	53.68	75.987	.451	.350	.690
Миний ажлын байрны нөхцөл таатай	53.66	75.913	.450	.375	.690
Миний хийдэг ажил утга учиртай	54.67	80.466	.369	.357	.704
Би хамт ажиллагсадтайгаа таатай харилцаатай	52.60	82.005	.307	.152	.714

(Ursachi G., 2015) нарын үзсэнээр Кронбах альфа коэффициент нь 0.6-0.7 хооронд байх нь сайн биш ч гэсэн, хангалттай үзүүлэлт юм. Уг судалгааг 95 хувийн ач холбогдлын түвшинд хийгдэж байгаа гэж үзэн Кронбах альфа нь 0.6-с дээш байхаар авч үзэж болно. Хэрэв судалгааг 99 хувийн ач холбогдлын түвшинд хийж байгаа тохиолдолд Кронбах альфа нь 0.7 эсвэл 0.8-с дээш байхаар сонговол зүйтэй. Энэ тохиолдолд туршилтын судалгаанд (pilot study) их ач холбогдол өгөх шаардлагатай ба асуултуудын ойлгомжтой байдал, үг найруулга зэрэгт анхаарах хэрэгтэй.

3.1.2 Корреляцийн шинжилгээ

Хосын буюу хоёр үзүүлэлтийн Корреляцын шинжилгээ нь хоёр тоон үзүүлэлтийн хоорондын шугаман хамаарлын зэрэг буюу тэдний хоорондын уялдаа хамаарлын чанга сулыг тодорхойлоход чиглэгдэнэ. Нэг эх олонлогоос сонгогдсон ижил n хэмжээтэй хоёр тоон үзүүлэлтийн ажиглалтын утга өгөгдөхөд Корреляцын коэффициентыг дараахь томъёогоор тооцно. Үүнд:

$$r = r_{xy} = \frac{Cov(x, y)}{\sqrt{Var(x) \times Var(y)}}$$

$Cov(x, y)$ - x, y үзүүлэлтийн коварианс

$\text{Var}(y)$ -у үзүүлэлтийн вариаци.

Корреляцийн коэффициент нь -1 эс $+1$ хязгаарт орших ба $r > 0$ бол хоёр үзүүлэлт шууд, $r < 0$ бол урвуу хамааралтай болно.

Корреляцийн шинжилгээний үр дүнг дараагийн хуудсаар харуулсан болно. Ихэнх корреляцийн коэффициентууд 0.000 бүхий p утгатай байгаа нь ач холбогдлын 1 хувийн түвшинд корреляци хамааралтай гэдгийг нотолно.

Хүснэгт 6. Корреляцийн шинжилгээний үр дүн

Correlations

		Би ажилдаа дуртай	Манай байгууллага ажилчдаа санаа тавьдаг	Байгууллагын зүгээс намайг үргэлж дэмжиж байдаг	Миний хийдэг ажил сонирхолтой	Би хамт ажиллагсдаа дуртай	Миний ажлын байр надад таалагддаг	Би цалиндаа сэтгэл ханамжтай байдаг	Миний ажлын байрны нөхцөл таатай	Миний хийдэг ажил утга учиртай	Би хамт ажиллагсадтайгаа таатай харилцаатай
Би ажилдаа дуртай	Pearson Correlation	1	.254**	.337**	.301**	.402**	.291**	.259**	.356**	.553**	.211**
	Sig. (2-tailed)		.000	.000	.000	.000	.000	.000	.000	.000	.000
	N	437	426	417	412	425	409	421	419	423	417
Манай байгууллага ажилчдаа санаа тавьдаг	Pearson Correlation	.254**	1	.566**	.129**	.168**	.086	.110*	.117*	.073	.153**
	Sig. (2-tailed)	.000		.000	.008	.000	.080	.023	.016	.134	.002
	N	426	444	423	419	433	416	427	425	427	425

Байгууллагын зүгээс намайг үргэлж дэмжиж байдаг	Pearson Correlation	.337**	.566**	1	.227**	.138**	.085	.076	.128**	.069	.180**
	Sig. (2-tailed)	.000	.000		.000	.004	.086	.120	.009	.158	.000
	N	417	423	438	412	425	409	419	420	420	419
Миний хийдэг ажил сонирхолтой	Pearson Correlation	.301**	.129**	.227**	1	.243**	.115*	.140**	.147**	.284**	.144**
	Sig. (2-tailed)	.000	.008	.000		.000	.021	.004	.003	.000	.003
	N	412	419	412	434	422	404	417	417	418	414
Би хамт ажиллагсдадаа дуртай	Pearson Correlation	.402**	.168**	.138**	.243**	1	.101*	.143**	.129**	.121*	.355**
	Sig. (2-tailed)	.000	.000	.004	.000		.038	.003	.008	.012	.000
	N	425	433	425	422	447	420	429	428	431	428
Миний ажлын байр надад таалагддаг	Pearson Correlation	.291**	.086	.085	.115*	.101*	1	.536**	.501**	.291**	.135**
	Sig. (2-tailed)	.000	.080	.086	.021	.038		.000	.000	.000	.006
	N	409	416	409	404	420	430	412	412	413	412

Би цалиндаа сэтгэл ханамжтай байдаг	Pearson Correlation	.259**	.110*	.076	.140**	.143**	.536**	1	.482**	.210**	.117*
	Sig. (2-tailed)	.000	.023	.120	.004	.003	.000		.000	.000	.016
	N	421	427	419	417	429	412	441	426	424	422
Миний ажлын байрны нөхцөл таатай	Pearson Correlation	.356**	.117*	.128**	.147**	.129**	.501**	.482**	1	.231**	.120*
	Sig. (2-tailed)	.000	.016	.009	.003	.008	.000	.000		.000	.014
	N	419	425	420	417	428	412	426	443	424	421
Миний хийдэг ажил утга учиртай	Pearson Correlation	.553**	.073	.069	.284**	.121*	.291**	.210**	.231**	1	.095
	Sig. (2-tailed)	.000	.134	.158	.000	.012	.000	.000	.000		.050
	N	423	427	420	418	431	413	424	424	443	424
Би хамт ажиллагсадтай гаа таатай харилцаатай	Pearson Correlation	.211**	.153**	.180**	.144**	.355**	.135**	.117*	.120*	.095	1
	Sig. (2-tailed)	.000	.002	.000	.003	.000	.006	.016	.014	.050	
	N	417	425	419	414	428	412	422	421	424	440

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Корреляцын шинжилгээний Корреляцын коэффициентын утгаас хамааруулан түүний зэргийг дараах байдлаар ангилна. Үүнд:

1. $0 \leq |r| < 0.5$ - Сул
2. $0.5 \leq |r| < 0.75$ - Мэдэгдэхүйц
3. $0.75 \leq |r| < 0.9$ - Нягт
4. $0.9 \leq |r| < 1$ - Хүчтэй
5. $|r| = 1$ - Төгс буюу функцийн хамааралтай гэж тус тус ангилах бөгөөд үүнээс тэдгээрийн хоорондын уялдаа хамаарал ямар байгааг харуулна.

Үр дүнгээс харвал ажилдаа дуртай, сэтгэл ханамжтай байх байдал нь ажлаа утга учиртай гэдэгт итгэх итгэлтэй мэдэгдэхүйц хамааралтай байдаг байна. Бусад хувьсагчид 0.5-с бага коэффициенттой тул сул хамааралтай гэж дүгнэнэ.

3.1.3 Регрессийн шинжилгээ

Олон хүчин зүйлийн шугаман регрессийн шинжилгээг хийхэд дараах таамаглалыг дэвшүүлдэг:

- Таамагласан утгын буюу загварын алдаа нь түүврийн шинж чанараас үл хамаарна
- Загварын алдаа нь нормал тархалттай байна
- Загварын алдаа нь хомоскедастик буюу тогтмол вариацитай байна
- Бүх хувьсагчдын хамаарал нь шугаман байна

Эдгээр таамаглалуудыг бүгдийг нь регрессийн шинжилгээний хүрээнд шалгах боломжтой байдаг. Эхлээд регрессийн шинжилгээг хийх бөгөөд таамаглалуудыг няцаагдах/дэмжигдэх эсэхийг шалгана.

Өмнө нь тодорхойлох статистикийн дагуу түүврийн 297 өгөгдлийг шинжлэх боломжтой байсан. Юуны түүрүүнд корреляцийн шинжилгээгээр хувьсагчдын хоорондын хамаарлыг тодорхойлох ба дараагаар шаталсан регрессийн шинжилгээг хийнэ.

Уг өгөгдлийн бааз нь 9 тайлбарлагч хувьсагчийг агуулна. Шаталсан регрессийн шинжилгээг хийхдээ:

- Хамаарах хувьсагчийг хамгийн сайн тайлбарлах буй тайлбарлах хувьсагчдыг сонгох, үүнд ач холбогдлын түвшин нь 0.05-с ихгүй байна
- Загварын бүх тайлбарлах хувьсагчдыг магадлалын утгыг тодорхойлох. Загварын тайлбарлах хувьсагчдын магадлалын утга 0.1-с их бол үнэлгээнээс хасах
- Бүх процессыг дахин, дахин давтаж загвар нь дараах шаардлагыг хангасан байна: 1) бүх ач холбогдолтой хувьсагчдыг сонгох, 2) загварт ач холбогдолгүй хувьсагч байхгүй байх.

Хүснэгт 7. Регрессийн үнэлгээний хураангуй

Model Summary ^e									
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	Change Statistics			Sig. F Change
						F Change	df1	df2	
1	.553 ^a	.306	.304	.740	.306	177.183	1	402	.000
2	.648 ^b	.420	.417	.677	.114	78.885	1	401	.000
3	.697 ^c	.486	.482	.638	.066	51.627	1	400	.000
4	.719 ^d	.517	.512	.620	.030	25.070	1	399	.000

- a. Predictors: (Constant), Миний хийдэг ажил утга учиртай
- b. Predictors: (Constant), Миний хийдэг ажил утга учиртай, Би хамт ажиллагсдадаа дуртай
- c. Predictors: (Constant), Миний хийдэг ажил утга учиртай, Би хамт ажиллагсдадаа дуртай, Байгууллагын зүгээс намайг үргэлж дэмжиж байдаг
- d. Predictors: (Constant), Миний хийдэг ажил утга учиртай, Би хамт ажиллагсдадаа дуртай, Байгууллагын зүгээс намайг үргэлж дэмжиж байдаг, Миний ажлын байрны нөхцөл таатай
- e. Dependent Variable: Би ажилдаа дуртай

Шаталсан регрессийн шинжилгээний процедур тус бүрд хувьсагч нэмж үнэлгээ хийснээр загварын тайлбарлах чадвар нэмэгдэнэ.

- R – хамаарах хувьсагчийн бодит утга болон таамагласан утгын хоорондын Пирсоны корреляцийн коэффициент
- R² – хамаарах хувьсагчийн вариацийн хувь

Детерминацийн коэффициентын байж болох утгууд:

- 0,01-0,09 - хувьсагч ба үл хамаарах хувьсагчдын хамаарал сул, онолын хувьд хангалтгүй гэж үзнэ
- 0,09-0,49 - хувьсагчдын хамаарал дунд зэрэг
- 0,49-1,00 - хувьсагч ба үл хамаарах хувьсагчдын хамаарал хангалттай хүчтэй, регрессийн загварыг судалгаанд хэрэглэх онолын үндэслэлтэй гэж үзнэ

Аливаа загварыг сонгохдоо детерминацийн коэффициентийг максимумчлах нөхцөлийг харгалзах ба дээрх хүснэгтээс харвал Загвар 4 хамгийн тайлбарлах чадвар ихтэй нь байна.

Хүснэгт 8. Регрессийн ANOVA тест

		ANOVA ^a				
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	96.941	1	96.941	177.183	.000 ^b
	Residual	219.944	402	.547		
	Total	316.885	403			
2	Regression	133.096	2	66.548	145.198	.000 ^c
	Residual	183.789	401	.458		
	Total	316.885	403			
3	Regression	154.106	3	51.369	126.229	.000 ^d
	Residual	162.779	400	.407		
	Total	316.885	403			
4	Regression	163.729	4	40.932	106.636	.000 ^e

Residual	153.156	399	.384		
Total	316.885	403			

- a. Dependent Variable: Би ажилдаа дуртай
- b. Predictors: (Constant), Миний хийдэг ажил утга учиртай
- c. Predictors: (Constant), Миний хийдэг ажил утга учиртай, Би хамт ажиллагсдадаа дуртай
- d. Predictors: (Constant), Миний хийдэг ажил утга учиртай, Би хамт ажиллагсдадаа дуртай, Байгууллагын зүгээс намайг үргэлж дэмжиж байдаг
- e. Predictors: (Constant), Миний хийдэг ажил утга учиртай, Би хамт ажиллагсдадаа дуртай, Байгууллагын зүгээс намайг үргэлж дэмжиж байдаг, Миний ажлын байрны нөхцөл таатай

Хүснэгт 9. Регрессийн үнэлгээний коэффициентүүд

Model		Coefficients ^a				
		Unstandardized Coefficients		Standardized Coefficients Beta	t	Sig.
		B	Std. Error			
	(Constant)	5.214	.109		48.007	.000
1	Миний хийдэг ажил утга учиртай	.279	.021	.553	13.311	.000
	(Constant)	3.883	.180		21.598	.000
2	Миний хийдэг ажил утга учиртай	.258	.019	.512	13.364	.000
	Би хамт ажиллагсдадаа дуртай	.185	.021	.340	8.882	.000
	(Constant)	3.521	.177		19.921	.000
3	Миний хийдэг ажил утга учиртай	.251	.018	.498	13.780	.000
	Би хамт ажиллагсдадаа дуртай	.166	.020	.306	8.403	.000
	Байгууллагын зүгээс намайг үргэлж дэмжиж байдаг	.110	.015	.260	7.185	.000
	(Constant)	3.233	.181		17.857	.000
4	Миний хийдэг ажил утга учиртай	.232	.018	.460	12.782	.000
	Би хамт ажиллагсдадаа дуртай	.157	.019	.290	8.160	.000
	Байгууллагын зүгээс намайг үргэлж дэмжиж байдаг	.102	.015	.242	6.844	.000
	Миний ажлын байрны нөхцөл таатай	.083	.016	.181	5.007	.000

a. Dependent Variable: Би ажилдаа дуртай

Регрессийн коэффициентын үр дүнгээс харвал SPSS дахь үнэлгээ нь 4 шаттай байсан ба шат тус бүрд 1 хувьсагч нэмсэн байна. Детерминацийн коэффициентын утгын дагуу Загвар 4-ийг сонгосон тул уг загварын хүрээнд үр дүнгийн шинжилгээг хийнэ.

Эндээс тайлбарлах хувьсагчдыг дараах байдлаар томъёолж болно:

$$Y' = 3.233 + 0.232 * x_1 + 0.157 * x_2 + 0.102 * x_3 + 0.083 * x_4$$

Үүнд: Y – ажлын сэтгэл ханамж, x нь тайлбарлах хувьсагчид (хүснэгт дахь дарааллаар)

Үнэлгээнд ашигласан бүх хувьсагчдын шкал хэмжигдэхүүн нь ижил учраас стандартчилагдсан коэффициентын ашиглахаас илүү стандартчилагдаагүй коэффициентыг ашиглана.

Регрессийн коэффициентын найдвартай байдлын статистик үнэлгээг

$$t=b/S$$

Энд S – регрессийн коэффициентын дундаж алдаа.

Ажиглалтаас бодогдсон t-шинжүүрийн t_n -утгыг өгөгдсөн утгын түвшин- α , чөлөөний ($v=n-m-2$) зэрэгт харгалзах хүснэгтийн утга t-тэй харьцуулж үзнэ. Регрессийн загварын найдвартай байдал нь $t_n > t_{\text{критик}}$ – байгаагаар тодорхойлогдоно.

Хүснэгт 10. Үнэлгээнээс хассан хувьсагчид

Excluded Variables ^a						
Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics Tolerance
1	Манай байгууллага ажилчдадаа санаа тавьдаг	.214 ^b	5.319	.000	.257	.995
	Байгууллагын зүгээс намайг үргэлж дэмжиж байдаг	.300 ^b	7.716	.000	.360	.995
	Миний хийдэг ажил сонирхолтой	.157 ^b	3.671	.000	.180	.919
	Би хамт ажиллагсдадаа дуртай	.340 ^b	8.882	.000	.405	.985
	Миний ажлын байр надад таалагддаг	.141 ^b	3.295	.001	.162	.915
	Би цалиндаа сэтгэл ханамжтай байдаг	.149 ^b	3.565	.000	.175	.956

	Миний ажлын байрны нөхцөл таатай	.241 ^b	5.863	.000	.281	.947
	Би хамт ажиллагсадтайгаа таатай харилцаатай	.160 ^b	3.899	.000	.191	.991
2	Манай байгууллага ажилчдадаасанаа тавьдаг	.164 ^c	4.341	.000	.212	.969
	Байгууллагын зүгээс намайг үргэлж дэмжиж байдаг	.260 ^c	7.185	.000	.338	.978
	Миний хийдэг ажил сонирхолтой	.083 ^c	2.060	.040	.102	.875
	Миний ажлын байр надад таалагддаг	.117 ^c	2.976	.003	.147	.911
	Би цалиндаа сэтгэл ханамжтай байдаг	.109 ^c	2.811	.005	.139	.942
	Миний ажлын байрны нөхцөл таатай	.207 ^c	5.439	.000	.262	.936
	Би хамт ажиллагсадтайгаа таатай харилцаатай	.048 ^c	1.169	.243	.058	.871
3	Манай байгууллага ажилчдадаасанаа тавьдаг	.027 ^d	.627	.531	.031	.671
	Миний хийдэг ажил сонирхолтой	.031 ^d	.798	.425	.040	.842
	Миний ажлын байр надад таалагддаг	.102 ^d	2.726	.007	.135	.907
	Би цалиндаа сэтгэл ханамжтай байдаг	.097 ^d	2.636	.009	.131	.940
	Миний ажлын байрны нөхцөл таатай	.181 ^d	5.007	.000	.243	.926
	Би хамт ажиллагсадтайгаа таатай харилцаатай	.010 ^d	.249	.804	.012	.854
4	Манай байгууллага ажилчдадаа санаа тавьдаг	.019 ^e	.457	.648	.023	.670

Миний хийдэг ажил сонирхолтой	.022 ^e	.583	.560	.029	.840
Миний ажлын байр надад таалагддаг	.023 ^e	.548	.584	.027	.716
Би цалиндаа сэтгэл ханамжтай байдаг	.020 ^e	.509	.611	.025	.752
Би хамт ажиллагсадтайгаа таатай харилцаатай	-.001 ^e	-.023	.982	-.001	.852

a. Dependent Variable: Би ажилдаа дуртай

b. Predictors in the Model: (Constant), Миний хийдэг ажил утга учиртай

c. Predictors in the Model: (Constant), Миний хийдэг ажил утга учиртай, Би хамт ажиллагсдадаа дуртай

d. Predictors in the Model: (Constant), Миний хийдэг ажил утга учиртай, Би хамт ажиллагсдадаа дуртай, Байгууллагын зүгээс намайг үргэлж дэмжиж байдаг

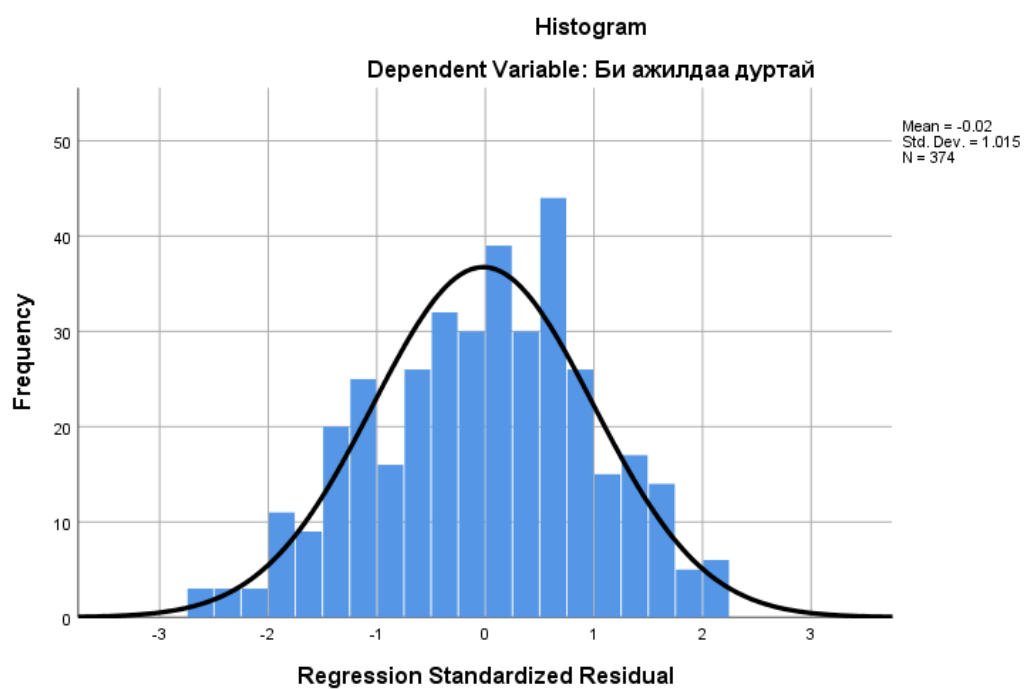
e. Predictors in the Model: (Constant), Миний хийдэг ажил утга учиртай, Би хамт ажиллагсдадаа дуртай, Байгууллагын зүгээс намайг үргэлж дэмжиж байдаг, Миний ажлын байрны нөхцөл таатай

Хүснэгт 11. Загварын үлдэгдлийн статистик

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	4.99	8.53	6.57	.639	391
Residual	-1.684	1.324	-.012	.629	374
Std. Predicted Value	-2.484	3.064	-.010	1.003	391
Std. Residual	-2.717	2.138	-.020	1.015	374

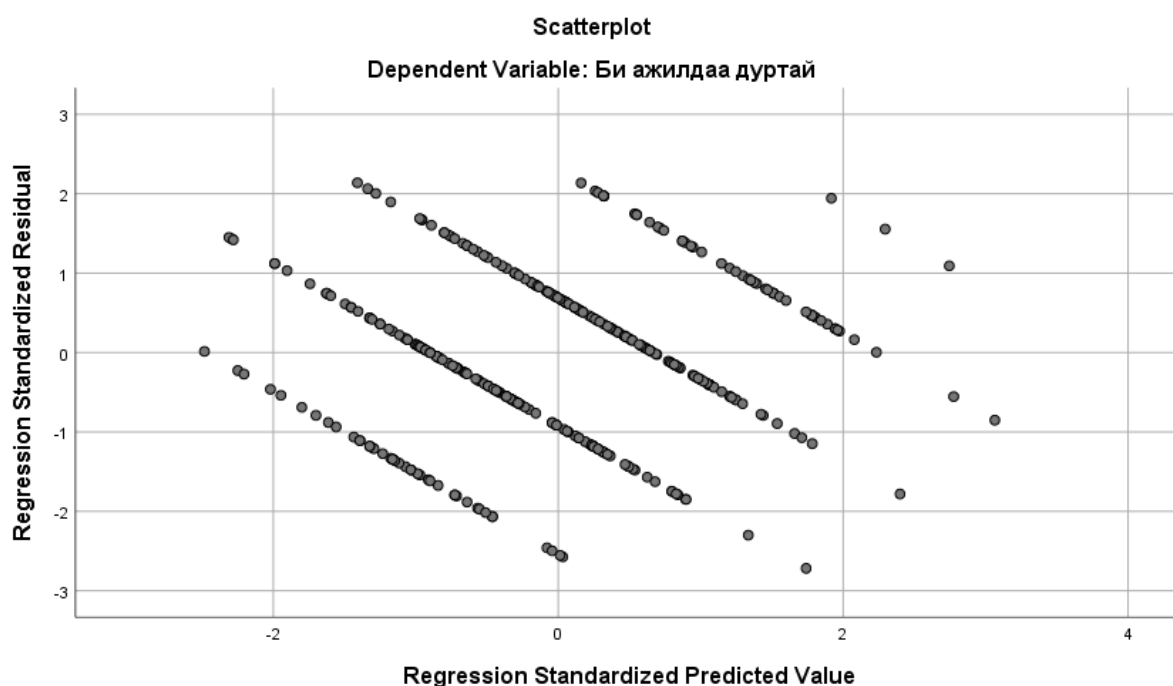
a. Dependent Variable: Би ажилдаа дуртай

Зураг 1. Регрессийн үлдэгдлийн гистограм



Регрессийн шинжилгээний гол таамаглалуудын нэг бол загварын үлдэгдлүүд нормал тархалттай байхад оршино. Гистограмын үр дүнгээс харвал үр дүнгүүд ерөнхийдөө нормал тархалтын муруйтай давхцаж байна. Иймд загварын үлдэгдлийн нормал тархалтын таамаглалыг дэмжих шаардлагатай.

Зураг 2. Регрессийн үлдэгдлийн тархалт



Регрессийн үнэлгээний үлдэгдлийн дисперси нь хэт их байх эсвэл хэт бага байх нь загварын алдааг илэрхийлнэ. Хомоскедастик нөхцөлийг дараах зүйлүүд тогтворгүй болгодог: огцом өндөр, бага үзүүлэлт, шаардлага хангаагүй өгөгдлүүд г.м.

Дээрх зургаас харвал үлдэгдлүүд нь ерөнхийдөө жигд бус тархсан байгаа бөгөөд иймд хомоскедастик нөхцөл биелж байгаа гэж үзэн, шинжилгээний өгөгдлийг цааш нь боловсруулах шаардлагатай гэж үздэг.

3.1.4 Үр дүнгийн шинжилгээ

Шаталсан регрессийн үр дүнг тайлбарлах нэгдсэн аргачлал байдаггүй тул регрессийн шинжилгээний үндсэн аргачлалуудын үр дүнгээр тайлбарлана:

- Тайлбарлах статистикийн үр дүнгүүд
- Хувьсагчдын хоорондын корреляцийн шинжилгээ
- Регрессийн үнэлгээний хураангуй хүснэгт
- Регрессийн үнэлгээний дэлгэрэнгүй хүснэгт

ДҮГНЭЛТ

Энэхүү судалгааны ажлаар шаталсан регрессийн аргачлалын оновчтой хэрэглээний цар хүрээ, арга зүйг нь судлах зорилго тавьсан. Зорилгынхоо хүрээнд шаталсан регрессийн арга, аргачлалыг судалж түүний хэрэглээг харууллаа. Ингэхдээ эхлээд асуулгын найдвартай байдлыг тодорхойлсон ба үүний дараагаар хувьсагчдын тодорхойлох статистикийн үзүүлэлтүүдийг тооцож корреляцийн шинжилгээ хийсэн юм. Дараагаар нь шаталсан регрессийн шинжилгээ болон загварын үлдэгдлийн шинжилгээ хийн үнэлэлт дүгнэлт хийлээ.

Шаталсан регрессийн шинжилгээгээ хийхдээ таамагласан утгын буюу загварын алдаа нь түүврийн шинж чанараас үл хамаарна, загварын алдаа нь нормал тархалттай байна, загварын алдаа нь хомоскедастик буюу тогтмол вариацитай байна, бүх хувьсагчдын хамаарал нь шугаман байна гэх таамаглалуудыг дэвшүүлэн шалгасан. Статистикийн дагуу түүврийн 297 өгөгдлийг шинжлэх боломжтой байсан ба корреляцийн шинжилгээгээр хувьсагчид хоорондын хамаарлыг тодорхойлж дараагаар нь шаталсан регрессийн шинжилгээгээ хийсэн. Шаталсан регрессийн шинжилгээг хийхдээ хамаарах хувьсагчийг хамгийн сайн тайлбарлаж буй тайлбарлах хувьсагчдыг сонгож загварын бүх тайлбарлах хувьсагчдын магадлалын утгыг тодорхойлон шаталсан регрессийн шинжилгээний процедур тус бүрд хувьсагч нэмж үнэлгээ хийсэн бөгөөд аливаа загварыг сонгохдоо детерминацийн коэффициентыг максимумчлах нөхцөлийг харгалзан үзэхэд загвар 4 хамгийн тайлбарлах чадвар ихтэй байсан тул тус загварт үр дүнгийн шинжилгээ хийсэн юм. Үр дүнг тайлбарлах нэгдсэн аргачлал байдаггүй тул регрессийн шинжилгээний үндсэн аргачлалуудын үр дүнгээр тайлбарлалаа.

Ерөнхийдөө шаталсан регресс нь тайлбарлагч хувьсагчдыг автоматчилагдсан алгоритмын дагуу сонгож, регрессийн загварыг оновчлох статистикийн аргачлалуудын нэг бөгөөд ихэвчлэн их хэмжээний өгөгдөл тандах судалгааны хэрэгслийн хувьд ашигладаг байна. Хэдийгээр уг аргачлал нь олон шүүмжийг дагуулдаг ч одоог хүртэл хэрэглэгдсээр байна.

НОМ ЗУЙ

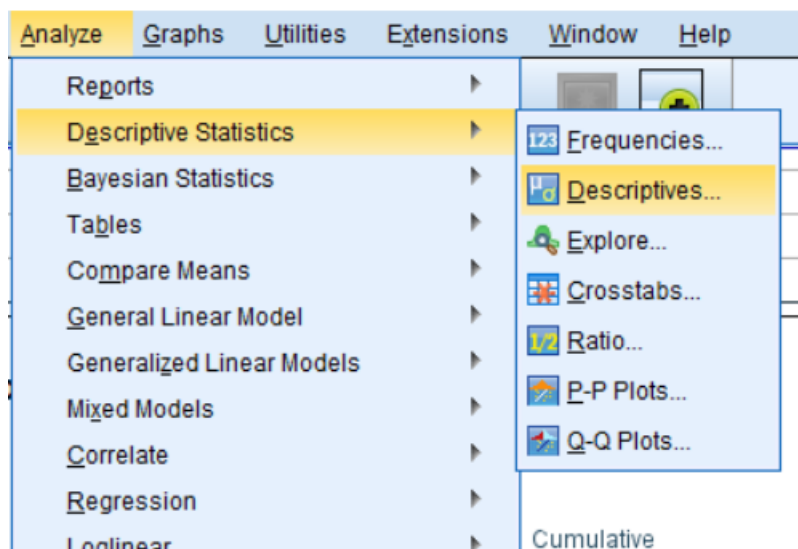
- Agresti, A. & Franklin, C. (2014). *Statistics. The Art & Science of Learning from Data*. Essex: Pearson Education Limited.
- Babyak MA. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosom Med*. 2004;66:411–21.
- Berry, W.D. (1993). *Understanding Regression Assumptions*. Newbury Park, CA:
- Castle JL, Fawcett NWP, Hendry DF. Evaluating automatic model selection, Technical Report 474. Oxford: Department of Economics, University of Oxford; 2010.
- Efroymson MA. Multiple regression analysis. In: Ralston A, Wilf HS, editors. *Mathematical methods for digital computers*. New York: Wiley; 1960.
- Flom PL, Cassell DL. Stopping stepwise: why stepwise and similar selection methods are bad, and what you should use. In: *NESUG 2007 proceedings*. 2007.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2013). *Multivariate data analysis: Advanced diagnostics for multiple regression [Online supplement]*. Retrieved from http://www.mvstats.com/Downloads/Supplements/Advanced_Regression_Diagnostics.pdf
- Hair, J.F., Black, W.C., Babin, B.J. et al (2006). *Multivariate Data Analysis*. New Jersey: Pearson Prentice Hall.
- Harrell FE Jr. *Regression modeling strategies: with applications to linear models, logistic regression and survival analysis*. New York: Springer; 2001.
- Hendry DF, Krolzig HM. *Automatic econometric model selection*. London: Timberlake Consultants Press; 2001.
- Howell, D.C. (2002). *Statistical Methods for Psychology (5th ed.)*. Pacific Grove CA: Duxbury.
- Huberty CJ. Problems with stepwise methods—better alternatives. In: Thompson B, editor. *Advances in social science methodology*, vol. 1. Greenwich: JAI Press; 1989.
- Hurvich CM, Tsai CL. The impact of model selection on inference in linear regression. *Am Stat*. 1990;44(3):214–7.
- IBM (n.d.). *Collinearity diagnostics*. from https://www.ibm.com/support/knowledgecenter/en/SSLVMB_23.0.0/spss/tutorials/reg_cars_collin_01.html

- Marascuilo LA, Serlin RC. Statistical methods for the social and behavioral sciences. New York: W. H. Freeman; 1988.
- Nicol, A.M. & Pexman, P.M. (2010). Presenting Your Findings. A Practical Guide for Creating Tables. Washington: APA.
- Sage. Field, A. (2013). Discovering Statistics with IBM SPSS Newbury Park, CA: Sage.
- Snee, R. D. (1983). Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. *Journal of Quality Technology*, 15, 149-153. doi:10.1080/00224065.1983.11978865
- Stevens, J. (2002). Applied multivariate statistics for the social sciences. Mahway, NJ: Lawrence Erlbaum Associates.
- Thompson B. Stepwise regression and stepwise discriminant analysis need not apply here: a guidelines editorial. *Educ Psychol Meas.* 1995;55:525–34.
- Thompson B. Why won't stepwise methods die? *Meas Eval Couns Dev.* 1989;21(4):146–8.
- Vlachopoulou M, Ferryman TA, Zhou N, Tong J. A stepwise regression method for forecasting net interchange schedule. <https://doi.org/10.1109/pesmg.2013.6672763>. 2013.
- Walter S, Tiemeier H. Variable selection: current practice in epidemiological studies. *Eur J Epidemiol.* 2009;24(12):733–6
- Whittingham MJ, Stephens PA, Bradbury RB, Freckleton RP. Why do we still use stepwise modelling in ecology and behaviour? *J Anim Ecol.* 2006;75(5):1182–9.

ХАВСРАЛТ

Хавсралт 1. Тодорхойлох статистикийн шинжилгээ хийх аргачлал.

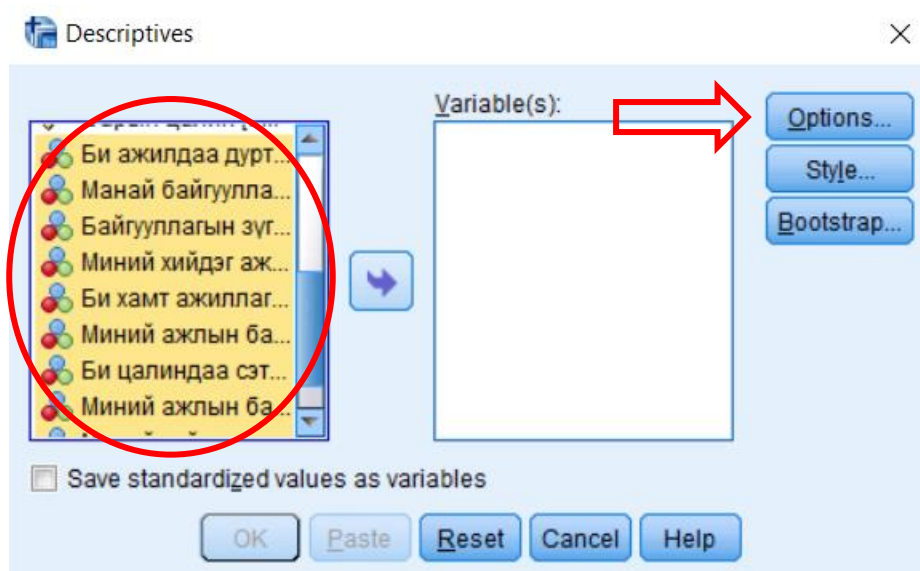
Зураг 3. Тодорхойлох статистикийн шинжилгээ хийх аргачлал. Алхам 1



Уг тест нь программын дараах цэсэнд байна:

Analyse > Descriptive Statistics > Descriptives

Зураг 4. Тодорхойлох статистикийн шинжилгээ хийх аргачлал. Алхам 2



Хувьсагчдын жагсаалтаас судалгааны үндсэн асуултууд болох хувьсагчдыг сонгож, идэвхжүүлсний дараагаар тохируулгын цэс рүү орно.

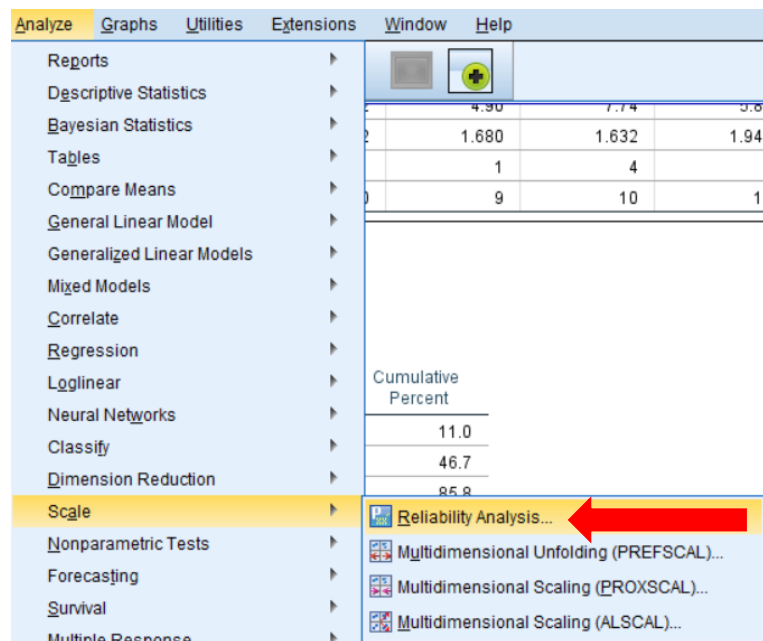
Зураг 5. Тодорхойлох статистикийн шинжилгээ хийх аргачлал. Алхам 3



Тохируулгын цэс хэсгээс тодорхойлох статистикийн үндсэн сонголтуудыг хийнэ: ерөнхий дундаж, стандарт хазайлт, доод ба дээд утга. Сонголтуудыг хийсний дараа тодорхойлох статистикийн тооцооллын цонх гарч ирнэ.

Хавсралт 2. Найдвартай байдлын шинжилгээ хийх үе шатууд.

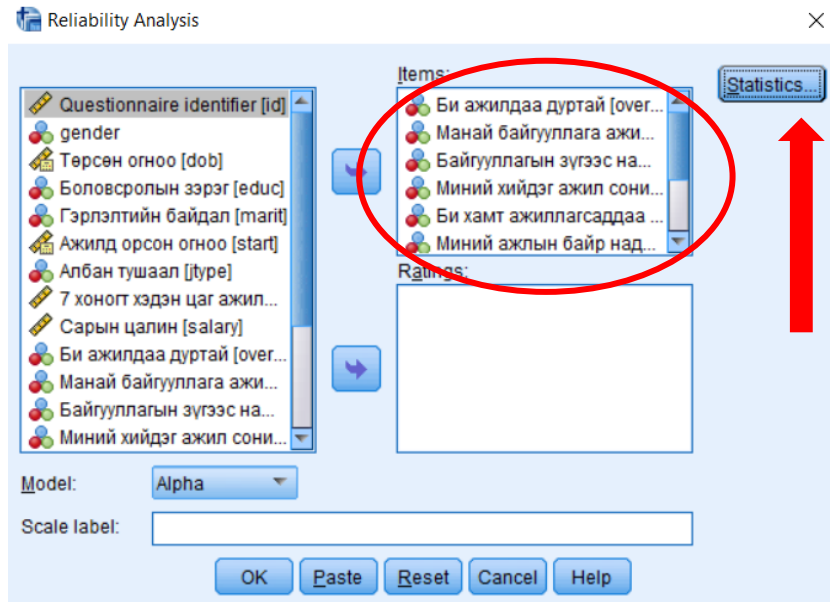
Зураг 6. Найдвартай байдлын шинжилгээ хийх үе шатууд. Алхам 1



Найдвартай байдлын шинжилгээг хийхдээ дараах цэсээс сонголт хийнэ.

Analyze > Scale > Reliability Analysis

Зураг 7. Найдвартай байдлын шинжилгээ хийх үе шатууд. Алхам 2



Хувьсагчдын жагсаалтаас судалгааны үндсэн асуултууд болох хувьсагчдыг сонгож, идэвхжүүлсний дараагаар тохируулгын цэс рүү орно.

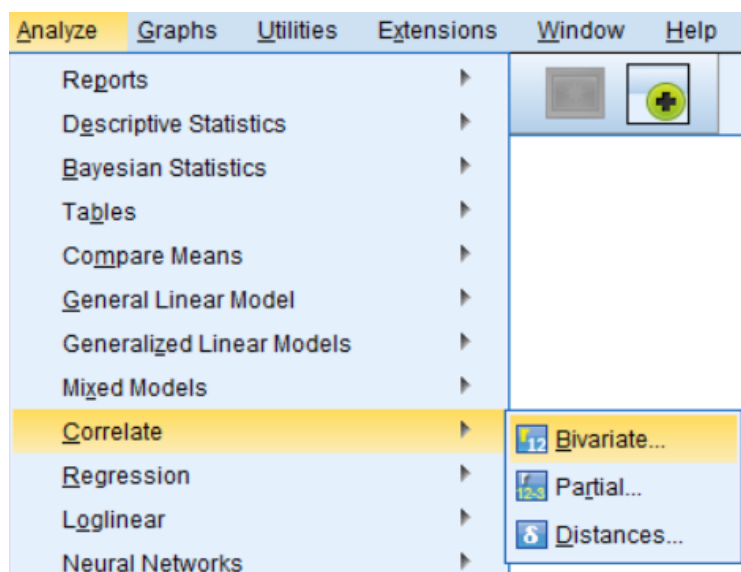
Зураг 8. Найдвартай байдлын шинжилгээ хийх үе шатууд. Алхам 3



Тохируулгын цэсээс дээрх зурагт харуулсанчлан сонголтуудыг хийнэ.

Хавсралт 3. Корреляцийн шинжилгээ хийх үе шатууд.

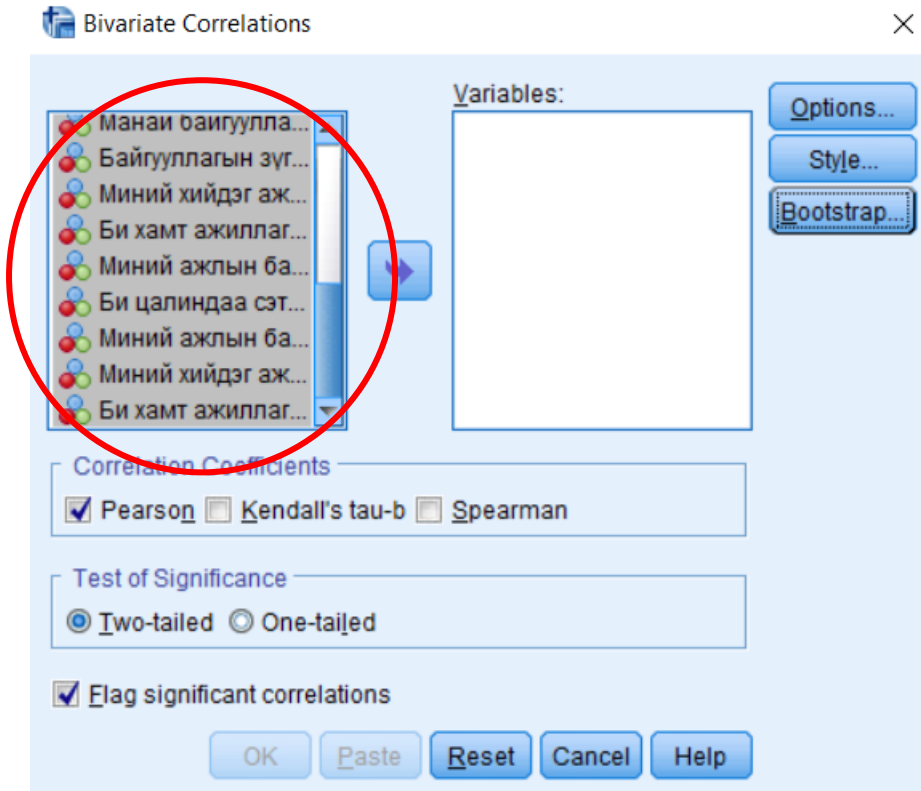
Зураг 9. Корреляцийн шинжилгээ хийх үе шатууд. Алхам 1



Корреляцийн шинжилгээг хийхдээ дараах цэсээс сонголт хийнэ:

Analyze > Correlate > Bivariate

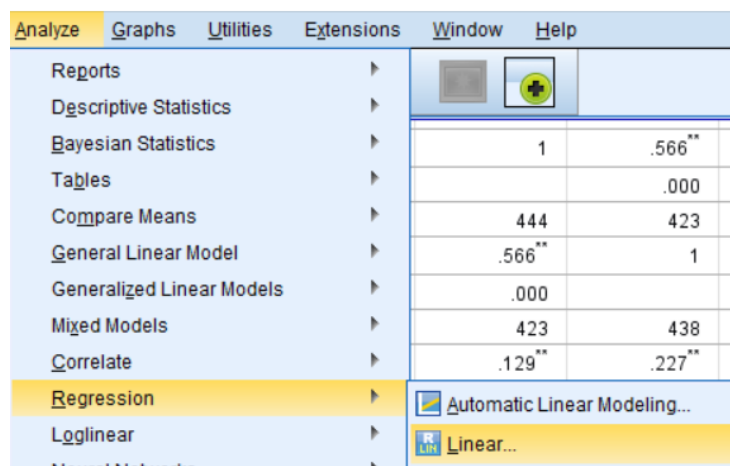
Зураг 10. Корреляцийн шинжилгээ хийх үе шатууд. Алхам 2



Корреляцийн шинжилгээг хийхдээ шаардлагатай хувьсагчдыг сонгоход хангалттай байна.

Хавсралт 4. Регрессийн шинжилгээ хийх үе шатууд

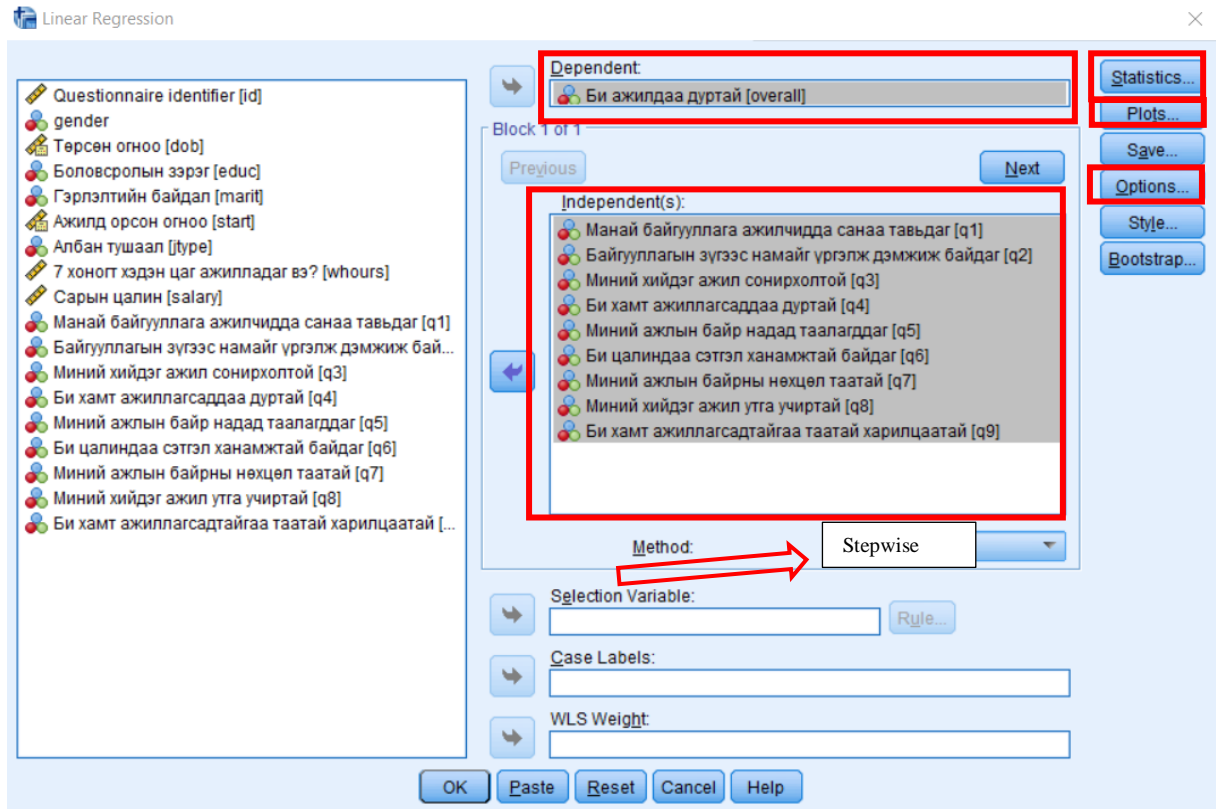
Зураг 11. Регрессийн шинжилгээ хийх үе шатууд. Алхам 1



Шаталсан регрессийн шинжилгээг хийхдээ энгийн шугаман регрессийн шинжилгээ хийх аргачлалыг ашиглана. Үүнд дараах цэсээс сонголт хийнэ:

Analyze > Regression > Linear

Зураг 12. Регрессийн шинжилгээ хийх үе шатууд. Алхам 2



Dependent нүдэнд хамаарах хувьсагчийг оруулах ба Independent гэсэн цонхонд үл хамаарах гэж үзсэн хувьсагчдыг оруулна. Энэ цонхонд Statistics, Plots болон Options гэсэн 3 товчийг дарж, тус бүрд нь харгалзах тохируулгыг хийнэ. Мөн Method буюу аргачлалыг Stepwise болгоно.

Method – ийн доош харсан сумыг дарахад регрессийн тэгшитгэлийн үзүүлэлтүүдийг тооцоолох 5 төрлийн алгоритмын аргыг сонгох боломж гарна. Үүнд:

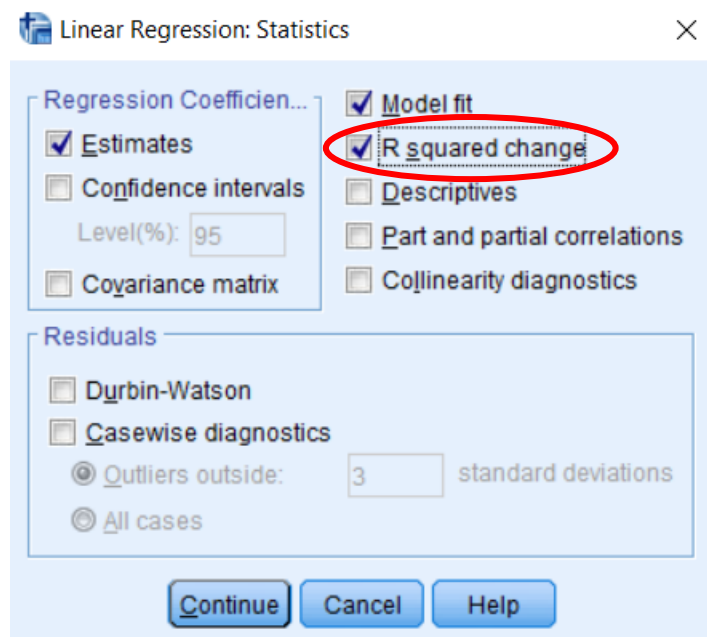
- Enter- бүх үл хамаарах хувьсагчдыг оролцуулан шууд нэг удаа бодож гаргасан тэгшитгэлийн үзүүлэлтүүдийг гаргадаг
- Stepwise – алхам алхамаар сайжруулахдаа нэмэх (нэг үл хамаарах хувьсагч эхэлж авах замаар) ба хасах (бүх үл хамаарах хувьсагчийг оруулаад) гэж 2 янзаар гүйцэтгэнэ. Нэмэхдээ тухайн тэгшитгэлд ороогүй хамгийн бага Fмагдлалтай үл хамаарах хувьсагчийг сонгоно. Хасахдаа тухайн тухайн тэгшитгэлд орсон хангалттай их Fмагдлалтай үл хамаарах хувьсагчийг сонгоно.

- Remove – хамаарах хувьсагчдыг тэгшитгэлийн чанарыг сайжруулах зорилгоор алхам бүрт нэгийг хасах замаар гүйцэтгэнэ.
- Backward- үл хамаарах бүх хувьсагчид тэгшитгэлд эхэлж ороод хамгийн бага хэсгийн корреляцтай үл хамаарах хувьсагчийг эхлэн хасч сүүлийн хасагдсанаас бага хэсгийн корреляцтай үл хамаарах хувьсагч гарч ирэхгүй болтол үргэлжлүүлнэ.
- Forward- Энд эхлээд хамаарах хувьсагчтай хамгийн их корреляцтай (-,+) үл хамаарах хувьсагч орох бөгөөд дараа нь хамгийн их хэсгийн корреляцтай үл хамаарах хувьсагчийг нэмж сүүлийн хасагдсанаас их хэсгийн корреляцтай үл хамаарах хувьсагч гарч ирэхгүй болтол үргэлжлүүлнэ.

Selection variable-д судалгааг хязгаарлах хувьсагчийг дүрмийн (Rule) мар нэг үзүүлэлтийн утгыг тодорхой тоогоор хязгаарлах) хамт тодорхойлно. Case Labels-д үнэлгээний үр дүнгийн график гаргахад ямар хувьсагчтай харьцуулан гаргахыг заана.

WLS –хамгийн бага квадратын аргад ашиглах жигнэх “weight” хувьсагчийн нэр. Тэр нь регрессийн тэгшитгэлд ороогүй байх бөгөөд хасах “-”, тэг “0” утга авч байвал хасагдана.

Зураг 13. Регрессийн шинжилгээ хийх үе шатууд. Алхам 3



Регрессийн шинжилгээ хийхэд ашиглагддаг, загварын тайлбарлах чадварын гол шалгуур үзүүлэлт болох детерминацийн коэффициентийг Statistics цонхноос идэвхжүүлнэ.

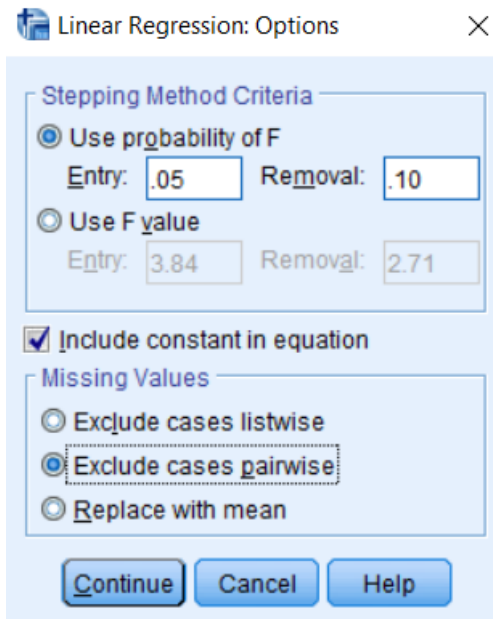
Statistics –т ямар статистик гаргаж авахыг сонирхож байгаагаа тодорхойлох:

- Regression Coefficients- засварлагдсан, итгэмжлэгдэх, ковариацийн гэсэн 3 боломжоор коэффициентуудыг үзүүлнэ.
- Model fit- регрессийн тэгшитгэлд ямар үзүүлэлтүүдийн тооцоог харуулах (R , R квадрат, засварлагдсан R квадрат, стандарт алдаа. Мөн ANOVA (analysis of variance) -вариацийн шинжилгээний чөлөөний зэрэг, квадратуудын нийлбэр, дундаж квадрат, F - утга, F -н ажиглалтын магадлал зэргийг тооцсон утгыг үзүүлнэ.
- R squared change- регрессийн үед үзүүлэлтийг тэгшитгэлээс нэмэх хасах зорилгоор R -квадратыг бодно. Хэрэв R квадрат ихэсч байхад дээрх үйл ажиллагаа хийгдэх бөгөөд тэр үзүүлэлтийг сайн предиктор гэж хэлнэ.
- Descriptives- дундаж, стандарт хазайлт, корреляцын матрицийг үзүүлнэ.
- Part and partial correlations- хэсгийн болон хагас хэсгийн Корреляцын коэффициентийг харуулах ба энэ нь -1 –ээс 1 -ийн хооронд байдаг. Энэ тэмдэг нь хамаарлын чиглэлийг заах бөгөөд түүний абсолют утга их байвал хамаарал хүчтэй байгааг харуулна.
- Collinearity diagnostics (or multicollinearity)- үл хамаарах хувьсагчдын шугаман хамаарлыг шалгах бөгөөд шугаман хамаарал илэрвэл арилгах шаардлагатай.

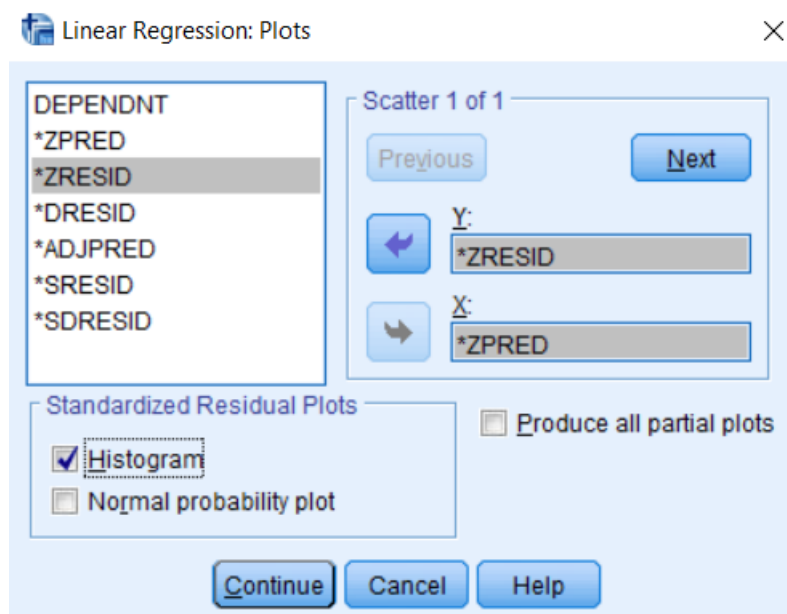
Residuals-д ажиглалтын болон онолын утгын зөрүүг харах хэлбэр.

- Displays the Durbin-Watson- нь үлдэгдлүүд хамаарах хувьсагчдын үечлэлтэйгээр хүчтэй корреляцтай байгааг шалгана.
- Casewise diagnostics- ажиглалт бүрээр (All cases) шалгаж өгөх бөгөөд хүсвэл стандарт хазайлтын хязгаар хэмжээг (стандарт нь 3) зааж зөвхөн түүнд хамаарах хазайлттайг шалгаж болно.

Зураг 14. Регрессийн шинжилгээ хийх үе шатууд. Алхам 4



Зураг 15. Регрессийн шинжилгээ хийх үе шатууд. Алхам 5



Plots-оос графикуудын хэлбэрийг сонгоно. (DEPENDNT)-Y хамаарах хувьсагчийг зааж дараагаар нь X-хамаарах хувьсагч болон үлдэгдэл үзүүлэлтүүдийг заана: Стандартчилагдсан хамаарах хувьсагчид (*ZPRED), Стандартчилагдсан үлдэгдлүүд (*ZRESID), хасагдсан үлдэгдлүүд (*DRESID), Засварлагдсан хамаарах хувьсагчид (*ADJPRED), Стьюдентын үлдэгдлүүд (*SRESID), стандартчилагдаж хасагдсан үлдэгдлүүд (*SDRESID) тус тус болно.