

ХҮН АМЫН МЭДЭЭЛЛИЙН ЧАНАРЫГ ҮНЭЛЭХ АРГА ЗҮЙН ЗАРИМ АСУУДАЛ

*Б. Мөнхжаргал, магистр, Эдийн засгийн тэнхимийн ахлах багш, Ниймгийн ухааны салбар,
Шинжлэх ухааны сургууль, Монгол Улсын Их Сургууль*

ОРШИЛ

Хүн ам зүйчдийн нэг үндсэн үүрэг бол хүн ам зүйн үнэлгээ, тооцоог үнэн зөв гүйцэтгэх явдал билээ. Үүний тулд өндөр чанартай, эсвэл алдааг нь засварлаж, шулуутгаж болох мэдээлэлтэй байх шаардлагатай. Мэдээлэлд засварлалт хийх боломжгүй ч алдааг үнэлэх нь их чухал, учир нь алдааны үнэлгээ бол тухайн мэдээллийн хүн ам зүйн тооцоо, шинжилгээнд ашиглагдаж болох итгэмжлэгдсэн түвшинг харуулдаг.

Хүн амын тоо мэдээллүүдийн чанарыг үнэлж дүгнэхийн өмнө бид мэдээллийн алдааны дараахь хоёр төрлийг ялгаж ойлгох хэрэгтэй. Эдгээр нь *хамрагдалтын алдаа* ба *агуулгын алдаа* юм. **Хамрагдалтын алдаа** нь мэдээллийн систем дэх үзэгдэл буюу хүмүүсийн бүртгүүлэлтийн бүрэн байдалтай холбоотой бөгөөд хүн ам, тэдний дунд гарч буй үзэгдлийг¹ дутуу тоолж бүртгэх, эсвэл давхардуулан тоолох тохиолдлыг багтаана. Харин **агуулгын алдаа** нь мэдээллийн системд бүртгэгдсэн үзэгдэл, юмсын шинж чанарын үнэн зөв байдалтай холбоотой ба үүнд тухайлбал хүмүүс насны мэдээлэл өгөхдөө хий насаа нэмж хэлэх, цэрэг татлагаас зугтаах, тэтгэвэрт эрт гарах гэх мэт тодорхой зорилгоор насаа буруу мэдээлэх, эсвэл шилжих хөдөлгөөний бүх тохиолдол бүртгэгдсэн боловч хугацаа нь буруу бүртгэгдэх, байршил нь буруу бүртгэгдэх гэх мэт тохиолдлуудыг хамруулж ойлгоно. Үзэгдэл цаг хугацаандаа бүртгэгдэхгүй байх нь хамрагдалтын алдааг үүсгэдэг. Тухайлбал, 2010 онд төрсөн хүүхдийг 2011 онд бүртгүүлэх нь 2010 оны мэдээллийг дутуу хамралттай, 2011 оны мэдээллийг давхардсан хамралттай болгож, хоёуланд нь хамрагдалтын алдаа гарахад хүргэж байна.

Дэлхийн ихэнх улс орнуудад, ялангуяа аж үйлдвэржсэн өндөр хөгжилтэй орнуудад хүн ам, түүний өөрчлөлт хөдөлгөөнийг илэрхийлэх үзэгдлүүдийн тоо мэдээллийн бүртгэл нь албан ёсны буюу хуульчлагдсан байдаг бөгөөд хүн амын тооллого, иргэний бүртгэлийн систем, мөн томоохон түүвэр судалгаа зэрэг эх үүсвэрүүдээр тогтмол хийгдэж байдаг. Энэ нь мэдээллийн чанарт ихээхэн эерэг нөлөө үзүүлдэг. Гэтэл хүн амын өөрчлөлт, хөдөлгөөний талаарх мэдээллийн систем сайн хөгжөөгүй орнуудад эсрэгээрээ мэдээллийн чанар тийм ч хангалттай байдаггүй тул бид нийгэм, эдийн засгийн хөгжлийн аливаа төлөвлөлт хийх, бодлого, хөтөлбөр боловсруулах, хэрэгжүүлэх, үнэлэхдээ болон нарийвчилсан судалгаа шинжилгээ хийхдээ мэдээллийн чанарыг аль болох үнэлж, шаардлагатай тохиолдолд алдааг засварлаж шулуутган ашиглах нь зүйтэй юм.

Хүн амын холбогдолтой тоо, мэдээллийн чанарыг үнэлэх хоёр арга байдаг. Эдгээрийн нэг нь, **харьцуулах арга** /matching approach/ буюу статистикчид голлон хэрэглэдэг арга, нөгөө нь хүн ам зүйн тэнцлийн томьёо²нд үндэслэдэг **хүн ам зүйн арга** юм. Хүн амын холбогдолтой тоо мэдээллийн алдааны төрлүүд болон тэдгээрийг үнэлэх аргуудыг дараахь хүснэгтэд нэгтгэн харуулав.

Алдааны төрөл	Алдаа тодорхойлох аргууд	
	Харьцуулах (статистик)	Хүн ам зүйн
Хамрагдалтын	1	3
Агуулгын	2	4

Сэтгүүлийн энэхүү дугаарт хамрагдалтын болон агуулгын алдааг үнэлэх дээрх 2 төрлийн аргуудын талаар дэлгэрэнгүй авч үзье.

¹ Хүн амын үзэгдэл гэдэгт төрөлт, нас баралт, шилжин ирэлт, шилжин явалт, гэрлэлт, цуцлалт, хөдөлмөрийн нөөцөд шилжин орох, тэтгэвэрт гарах гэх зэрэг хүн амын хийдэг олон төрлийн үйлдэл, хөдөлгөөнийг хамруулан ойлгоно.

² Хоёр цаг хугацаа хоорондын хүн амын тооны өөрчлөлт нь тухайн хугацааны туршид явагдсан байгалийн цэвэр өсөлт болон механик цэвэр өсөлтийн нийлбэрээр тодорхойлогдоно гэдгийг харуулсан илэрхийлэл. Томьёолбол, (Тайлант үеийн хүн ам - Суурь үеийн хүн амын тоо) = (Төрөлт - Нас баралт) + (Шилжин ирэгчид - Шилжин явагчид)

1. ХАМРАГДАЛТЫН АЛДААГ ҮНЭЛЭХ СТАТИСТИК АРГА

Мэдээллийн хамрагдалтын алдааг үнэлэх гэдэг нь мэдээллийн системд бүртгэгдвэл зохих мэдээллүүдийн хичнээн хувь нь бодитойгоор бүртгэгдсэн бэ гэдгийг тогтоох явдал юм. Хүн амын холбогдолтой тоо мэдээллийн хамрагдалтын алдааг үнэлэх статистик арга нь нэг мэдээллийн эх үүсвэрт бүртгэгдсэн мэдээллийг өөр эх үүсвэрт бүртгэгдсэн ижил мэдээлэлтэй харьцуулах замаар хамрагдалтын түвшинг нь тооцоолдог арга юм. Тухайлбал, тодорхой хугацааны туршид иргэний бүртгэлийн системээр бүртгэгдсэн төрөлтийг түүвэр судалгаагаар эмэгтэйчүүдээс өнгөрсөн үеийн төрөлтийнх нь түүхийг асуух замаар цуглуулсан ижил цаг хугацааны үеийн төрөлтийн мэдээлэлтэй харьцуулж болно. Энэ арга 1960, 1970-аад оны үеэс хөгжиж буй орнуудын төрөлт, нас баралтын коэффициентуудыг үнэлэхэд ихээхэн дэвшил үзүүлсэн юм.

Энэхүү аргын логик нь Chandra-Sekar and Deming (1949) нарын “давхар бүртгэх үйлдэл”-ийн тухай сонгодог бүтээлд тэмдэглэгдсэн байдаг. Аргыг дараахь байдлаар тайлбарлая. Бидэнд иргэний бүртгэл, нөхөн үржихүйн насны эмэгтэйчүүдийн түүвэр судалгаа гэсэн хоёр эх үүсвэрээр цуглуулсан төрөлтийн талаарх мэдээлэл байна гэж үзье. Дээрх хоёр эх үүсвэрийн мэдээллийг харьцуулж, үр дүнг дараахь хүснэгтийн байдалтайгаар харуулж болно.

Төрөлтийн тоо		Түүвэр судалгаагаар	
		Бүртгэгдсэн	Бүртгэгдээгүй
Иргэний бүртгэлээр	Бүртгэгдсэн	100 (A)	50 (C)
	Бүртгэгдээгүй	20 (B)	(D)

Иргэний бүртгэлээр 150 төрөлт, түүвэр судалгаагаар 120 төрөлт бүртгэгдсэн; 100 тохиолдол нь хоёр мэдээллийн эх үүсвэрт хоёуланд нь бүртгэгдсэн байна. Нийтдээ 170 төрөлт (иргэний бүртгэлийн системээр бүртгэгдсэн 150, түүвэр судалгаагаар бүртгэгдсэн боловч иргэний бүртгэлд бүртгэгдээгүй 20) байна гэж хэлж болно. Энэ бол Chandra-Sekar & Deming нарын авч үзсэн Энэтхэгийн Иргэний Бүртгэлийн Системд тохиолдсон жишээ юм. Гэвч хоосон байгаа (D) нүдэнд төрөлтийн ямар нэг тоо буюу хоёр эх үүсвэрт хоёуланд нь

бүртгэгдээгүй тохиолдлын тоо байх ёстой. Энэ тоог тооцохдоо, иргэний бүртгэлд дутуу хамрагдах магадлал нь түүвэр судалгаанд дутуу хамрагдах магадлалтай хамааралгүй гэж төсөөлбөл:

$$\frac{A}{B} = \frac{C}{D} \text{ эндээс } D = \frac{C}{A} \cdot B = \frac{50}{100} \cdot 20 = 10$$

Энэ төсөөлөл нь бидний авч үзсэн хугацааны туршид нийт 180 төрөлт бүртгэгдэх ёстой байжээ гэсэн дүгнэлтэд хүргэж байна. Нөгөө талаас, түүвэр судалгааг харьцуулах үзүүлэлт болгон дээрх тооцоог хийж үзвэл, мэдээллийн хамрагдалтын түвшин $100/120=0.8333$ болно. Эндээс, нийт бүртгэгдэх ёстой төрөлтийн тоо нь $150/0.8333=180$.

Дээрх үйлдэл нэг эх үүсвэрт бүртгэгдэхгүй орхигдох магадлал нөгөө эх үүсвэрт бүртгэгдэхгүй орхигдох магадлал хоёрын хооронд ямар нэг корреляци-харилцан хамаарал байхгүй гэж төсөөлсөн. Энэ төсөөлөл бодит байдалд төдийлөн нийцдэггүй, учир нь практикт хоёр эх үүсвэрт бүртгэгдэхгүй орхигдох тохиолдлын тоо нь дундаж утгаасаа илүү байж болох тодорхой дэд бүлгүүдэд хуваагдсан мэдээллүүд элбэг тохиолддог. Доорх жишээнд бид дутуу хамрагдах магадлал нь ялгаатай хоёр тусдаа бүлгээр хүн амын мэдээллийг цуглуулсан гэж үзье.

1-р бүлэг:

Төрөлтийн тоо		Түүвэр судалгаагаар	
		Бүртгэгдсэн	Бүртгэгдээгүй
Иргэний бүртгэлээр	Бүртгэгдсэн	30	30
	Бүртгэгдээгүй	15	(15)

2-р бүлэг:

Төрөлтийн тоо		Түүвэр судалгаагаар	
		Бүртгэгдсэн	Бүртгэгдээгүй
Иргэний бүртгэлээр	Бүртгэгдсэн	70	20
	Бүртгэгдээгүй	5	(1.4)

Энд, (A), (B), (C) нүднүүдэд байгаа мэдээллүүдийн нийлбэр тоо нь бидний өмнө авч үзсэн жишээнийх гэдгийг тэмдэглэе. Гэхдээ дээр авч үзсэн хамааралгүй байх төсөөлөл 1 ба 2-р бүлэг тус бүрд мөн адил хүчинтэй гэж үзье. Тэгвэл хоёр мэдээллийн эх үүсвэрт зэрэг орхигдсон төрөлтийн нийт

тоо нь $(15+1.4)=16.4$ болж, 2 тусдаа бүлгээр мэдээлэл цуглуулахаас өмнө тооцогдсон 10 гэсэн утгаас илүү гарч байна. Энэхүү жишээ нь мэдээллийн эх үүсвэрүүдэд дутуу хамрагдах магадлалуудын хоорондын харилцан хамааралтай холбоотой “корреляцийн биас” хэмээх асуудлыг үлгэрлэн харуулсан жишээ юм. Өөрөөр хэлбэл, 2 жишээнд тооцоолсон дутуу хамрагдалтын түвшингүүдийн зөрүү нь коррелиацийн нөлөөгөөр үүссэн биас юм.

Онолын хувьд иймэрхүү төрлийн бэрхшээлийг шийдвэрлэх нэг арга бол, дутуу хамрагдах магадлалууд нь корреляцийн нөлөөнд орохгүй байж болох тийм дэд бүлгүүдэд хүн амыг хуваан судлах арга байж болох юм. Гэсэн ч үзэгдэл мэдээллийн ямар ч системд бүртгэгдэхгүй байх, эсвэл үзэгдлийн ямар нэг шинж тэмдэг нь корреляци үүсгэж болох нөхцлүүд ямагт оршсоор байдаг.

Нэмж хэлэхэд, корреляцийн биас нь зөв харьцуулах үзүүлэлт чухам аль нь болохыг тодорхойлоход бэрхшээл учруулдаг. Хоёр мэдээллийн эх үүсвэрт бүртгэгдсэн мэдээллүүд дотор харьцуулах үзүүлэлт болгон хэрэглэж болохуйц мэдээлэл байгаа эсэхийг тодорхойлох аргуудын тухай Ньюкомбе (1988)-д бичигдсэн байдаг. 1980 оны АНУ-ын хүн амын тооллогын үеэр, тухайн үеийн түүвэр судалгааны мэдээллийг харьцуулах үзүүлэлт болгон тооллогын мэдээллийг үнэлэхэд 9 хувийн дутуу хамралттай гэж гарсан. Гэтэл, тооллогын дутуу хамралт нь хичнээн их тодорхой бус байдалтай байсан ч гэсэн 2-оос илүү хувь бараг байдаггүйг бодолцвол, дээрх харьцаа хэт өндөр гэж тооцогдохоор байна. Иймээс, харьцуулах үзүүлэлтийг буруу сонгосон тохиолдолд энэ нь аливаа мэдээллийн хоёр эх үүсвэрт хоёуланд нь бүртгэгдэхгүй орхигдож болох магадлалын тооцоог хөөрөгддөг ба эх үүсвэрүүдэд бүрэн хамрагдах хамрагдалтын түвшингийн тооцоог багасгаж харуулдаг.

Төрөлт, нас баралтын мэдээллийг харьцуулах аргаар үнэлэх үед гарч ирдэг өөр нэг гол хүндрэл бол “хянах боломжгүй зүйлс”-ийн нөлөө юм. Энэ нь 2 системд бүртгэгдсэн мэдээлэл яг ижил цаг хугацаа юм уу, эсвэл ижил газар зүйн байршилд хамаарч чадахгүй үед үүсдэг биас буюу гадны нөлөө гэж болно. Эдгээр бэрхшээлүүдээс үүдэн төрөлт болон

нас баралтын мэдээллийн хамрагдалтын түвшинг үнэлэхэд харьцуулах аргыг хэрэглэх хандлага практикт харьцангуй бага байдаг. Гэхдээ энэ арга нь зарим оронд тооллогын хамрагдалтын түвшинг үнэлэхэд өнөөг хүртэл хэрэглэгдсээр байна.

2. АГУУЛГЫН АЛДААГ ҮНЭЛЭХ СТАТИСТИК АРГУУД

Агуулгын алдааг үнэлэх харьцуулалтын арга нь аливаа мэдээллийн эх үүсвэрээс авсан мэдээллийн агуулгыг нь шалгах замаар тухайн мэдээллийн эх үүсвэрүүдийн үнэн зөв байдлыг тогтоох, эсвэл хоёр мэдээллийн эх үүсвэрийн аль нэгнийх нь үнэн зөв байдлыг тогтооход чиглэгдсэн байдаг. Хоёр дахь зорилгын хувьд, хоёр мэдээллийн системийн аль нэг нь гарцаагүй үнэн зөв гэж батлагдсан тохиолдолд яригдана.

Нас бол хүн ам зүйн тооцоо, судалгаанд маш чухал үүрэг гүйцэтгэдэг учраас олонх хүн ам зүйчид насны мэдээллийн чанарын асуудлыг ихэд сонирхдог. Насны мэдээллийн үнэн зөвийг үнэлэх нэг энгийн арга бол хүн амын тооллогоор бүртгэгдсэн мэдээллийг тооллогоос богино хугацааны дараа явуулсан түүвэр судалгааны үр дүнтэй харьцуулан судлах арга юм. Дэлхийн янз бүрийн улс орнуудад ийм аргаар хийгдсэн судалгаануудыг харьцуулж үзэхэд зарим нийтлэг алдаанууд байгаа нь илэрсэн байна (Ewbank, 1981).

Хамгийн түгээмэл тохиолддог алдаануудын нэг бол насыг өсгөж хэлэх алдаа. 1966 онд Калдвеллийн (1966) Гана улсын хүүхдүүдийн дунд хийсэн насны мэдээллийн үнэн зөв байдлыг тогтоох судалгааны үр дүнгээс харахад нийт хүн амын 65 хувь нь сүүлчийн төрсөн өдрийн насаа үнэн зөв хэлсэн бол 9 хувь нь дутуу, 26 хувь нь өсгөж хэлсэн байна. Түүнчлэн Калдвелл нас ахих тусам насаа буруу мэдээлэх явдал өсч байгааг олж тогтоосон байна.

Нас баралтын насыг зөв тогтоохын тулд, Preston, Elo, Rozenwaike, Hill (1996) нар 1985, эсвэл 1980 онд нас барсан Африк гаралтай Америкчүүдийн нас баралтын гэрчилгээний мэдээллийг тэднийг хүүхэд болон залуу

настай байсан үед буюу 1900, 1910, 1920 онд явагдсан АНУ-ын хүн амын тооллогын мэдээлэлтэй харьцуулж үзжээ. Тэд мөн нас баралтын гэрчилгээний мэдээллийг нас нь илүү бодьтой бүртгэгдсэн байж болох Нийгмийн хамгааллын байгууллагын

нас баралтын бүртгэлийн мэдээлэлтэй ч харьцуулж үзсэн байна. Өөр хоорондоо ижил бус мэдээллийг хооронд нь харьцуулахдаа тэд тухайн хүний нэр, эцгийн нэр, эхийн нэр, төрсөн газар зэрэг үзүүлэлтүүдийг ашиглажээ.

Хүснэгт 1: Нийт хүн амын хувийн жин, насны бүлэг, насаа үнэн зөв мэдээлсэн байдлаар, Гана улсын иргэний бүртгэлийн систем, 1963 он.

Нас	Хэлсэн нас нь бодит наснаасаа дутуу байгаа (жилээр)			Хэлсэн нас нь бодит настайгаа тохирч байгаа	Хэлсэн нас нь бодит наснаасаа илүү байгаа (жилээр)		
	3	2	1		1	2	3
0	-	-	-	99	1	-	-
1	-	-	1	76	23	-	-
2	-	-	1	66	27	5	1
3	-	-	9	63	25	3	-
4	-	1	7	61	25	4	2
5	-	2	11	62	25	-	-
Бүх насныхан	1	1	7	65	22	3	1

Эх сурвалж: Caldwell, 1966.

Энэ судалгааны үр дүн нь нас баралтын гэрчилгээн дээр бичигдсэн нас барах үеийн наснууд ерөнхийдөө бага, залуу байгааг харуулсан. Гэтэл 95 ба түүнээс дээш насанд маш олон нас баралт бүртгэгдсэн байсан юм. Энэхүү логик зөрчилдөөн бол аль нэгэн насны интервалд бүртгэгдсэн нас баралтын тоо нь зөвхөн насаа буруу мэдээлсний цэвэр үр дагавар төдийгүй тухайн хүн амын үндсэн насны бүтэцтэй ч мөн холбоотой байдгаас гарч байгаа аж. Зөв насны бүтцийг ажиглавал нас баралтын нөлөөгөөр нас ахих тусам огцом буурсан хэлбэртэй байдаг. Иймд, хувь хүмүүсийн түвшинд насаа мэдээлэх үед бий болдог биас буюу аливаа нөлөөлөл нь нийт хүн амын мэдээллийг нэгтгэхэд заавал хувь хүний түвшин дэх шигээ ижил нөлөө үзүүлэх албагүй.

Мэдээж, харьцуулах судлал (matching study) нь зөвхөн насыг судлахад чиглэсэн байдаггүй бөгөөд ерөнхийдөө тооллогоор цуглуулагдсан бүх л үзүүлэлтүүдийг шинжлэхэд ашиглагдаж болно. Өөр хоорондоо ялгаатай хувьсагчуудын найдвартай/үнэн зөв байдлыг голдуу тооллогын дараахь түүвэр судалгаагаар харьцуулан судалдаг.

3. МЭДЭЭЛЛИЙН ЧАНАРЫГ ҮНЭЛЭХ ХҮН АМ ЗҮЙН АРГУУД

3.1. Нийцэмжтэй байдлын үнэлгээ

Мэдээллийн чанарыг шинжлэх хүн ам зүйн бүхий л аргууд нь хүн ам зүйн тооцооны тэнцэтгэлд үндэслэдэг. Жишээлбэл, үзэгдлүүдийн тоо хэмжээ үнэн зөв бүртгэгдсэн тохиолдолд уламжлалт хүн ам зүйн тэнцлийн томъёог ашиглан хоёр тооллогын хоорондох нийт хүн амын тоо хэмжээний өөрчлөлтийг дараахь тэнцэтгэлээр илэрхийлж болдог:

$$\Delta N = B - D + I - O \tag{1}$$

Энд, ΔN - 2 тооллогын хоорондох нийт хүн амын тооны өөрчлөлт,

B - 2 тооллогын хооронд бүртгэгдсэн төрөлтийн тоо,

D - 2 тооллогын хооронд бүртгэгдсэн нас баралтын тоо,

I - 2 тооллогын хооронд бүртгэгдсэн шилжин ирэлтийн тоо,

O - 2 тооллогын хооронд бүртгэгдсэн шилжин явалтын тоо.

Мөн энд, бусад үзэгдлүүдийн утга нь мэдэгдэж байх тохиолдолд аль нэг

үзэгдлийнх нь утгыг тооцон олж болдог. Харин мэдээллийн үнэн зөв байдлыг тогтоохын тулд уг тэнцэтгэлийг ашиглаж байгаа тохиолдолд тэнцэтгэл баланс барихгүй бол уг зөрүү нь мэдээллүүдээс ядаж нэг нь алдаатай, эсвэл үл нийцэх мэдээлэл байна гэдгийг харуулдаг.

Эхний тооллогын үед амьд байсан хүмүүсийн үелрэл³-ийн хувьд мэдээллийн нийцэмжтэй байдлыг шинжлэхэд мөн дээрх тэнцэтгэлийг ашиглаж болох ба энэ нь дараахь байдалтай томъёологдоно (энэ тохиолдолд В үзэгдэл тэгтэй тэнцүү болно):

$$\Delta N_c = -D_c + I_c - O_c \quad (2)$$

Энд, D_c , I_c , O_c нь тус тус хоёр тооллогын хоорондох с үелрэлийн нас барсан, шилжин ирсэн, шилжин явсан хүмүүсийн бодит тоо /илүүгүй, дутуугүй/.

(2) тэнцэтгэлийн дараахь хувилбар нь t ба $t+y$ үед хийгдсэн тооллогуудаар бүртгэгдсэн нас баралтын бүртгэлийн нийцэмжтэй байдлыг үнэлэхэд өргөнөөр хэрэглэгддэг (Condran et al., 1991):

$$R_c = \frac{N_c(t+y)}{N_c(t) - D_c + I_c - O_c} \quad (3)$$

Энд, $N_c(t)$ нь t үед үелрэлд тоологдсон хүн амын тоо

D_c нь тооллого хоорондын хугацаанд үелрэлээс нас барсан хүмүүсийн тоо

I_c нь тооллого хоорондын хугацаанд үелрэлд шилжин ирсэн хүмүүсийн тоо

O_c нь тооллого хоорондын хугацаанд үелрэлээс шилжин явсан хүмүүсийн тоо

(3) тэнцэтгэл дэх R_c нь хоёр дахь тооллогын үеийн: тоологдсон хүн амыг тооцоолсон хүн амын тоонд харьцуулсан харьцаа юм. R_c -ийн нэгээс хазайх хазайлт нь аль нэг, эсвэл бүх мэдээллийн эх үүсвэрүүд дэх насаа буруу мэдээлсэн байдал буюу хамрагдалтын алдаанаас шалтгаална. Ерөнхийдөө, насаа буруу мэдээлэх алдаа нь ялангуяа өндөр насныханы хувьд их

тохиолддог ба мэдээллийн хамрагдалт дутуу байхаасаа харьцангуй илүү анхаарал татахуйц асуудал юм. Мөн I_c , O_c -ийн талаарх мэдээлэл ч голдуу найдвартай бус байдаг. Учир нь хууль бусаар болон бичиг баримтгүй шилжин хөдөлгөөнд оролцогчид тооцооны найдвартай байдалд сөргөөр нөлөөлөх боломжтой. Кондран., нар (1991) янз бүрийн алдаанууд нийлээд хэрхэн R_c -ийн насны хэв маягт нөлөөлж байгааг судалж үзэхэд нас ахих тутам нас баралтын мэдээлэлд насыг өсгөж хэлэх алдаа нэмэгдсэн хандлагатай байжээ.

Эдгээр аргууд нь нийцэмжтэй байдлын үнэлгээнүүд юм. Нийцэмжтэй байдлын үнэлгээг үнэн зөв байдлын үнэлгээ (test of accuracy) болгон хувиргах 2 төрлийн боломж байдаг. Эхний боломж нь мэдээллийн эх үүсвэрүүдийн нэг буюу хэд хэдэн мэдээллийг бусдаасаа “давуу” гэж үзэн, түүнийг илүү бодьтой, бусад мэдээллийг нь үнэлэхэд хэрэглэж болно гэж найдах арга юм. Хоёр дахь нь загвар ашиглах арга. Эдгээр хоёр аргын талаар одоо авч үзье.

3.2 “Давуу” нэг буюу хэд хэдэн үзүүлэлтээр мэдээллийг үнэлэх

Иргэний бүртгэлийн систем сайн хөгжсөн орнуудад хүн амын тооллогын хамрагдалтыг үнэлэхдээ хэрэглэдэг гол стратеги нь төрөлт, нас баралт, шилжих хөдөлгөөний тоонууд үнэн зөв бөгөөд бүгд тооллогын бус эх үүсвэрээс цуглуулагдсан, мэдэгдэж байгаа буюу “давуу” талтай мэдээллүүд гэсэн төсөөлөл дээр үндэслэдэг. Эдгээр төрөлт, нас баралт, шилжих хөдөлгөөний тоонууд нь тооллого явагдсан хугацаан дахь төрөлтийн үелрэл бүрийн “бодит” хэмжээг тооцоход хэрэглэгддэг ба эдгээр тооцоог тооллогоор тоологдсон дүнтэй харьцуулж тооллогын хамрагдалтын алдааг үнэлдэг. Үелрэл дэх хүмүүсийн бодит тоог дараахь байдлаар тооцно:

$$\hat{N}_c = B_c - D_c + I_c - O_c \quad (4)$$

Энд, тэнцэтгэлийн баруун гар тал дахь хувьсагчууд нь бидний олох гэж буй үелрэлийн “бодит” хэмжээг тооцож байгаа ба тухайн үелрэл төрж эхэлснээс хойшхи өсөн нэмэгдсэн хэмжээг харуулна.

3 Үелрэл гэдэг нь судлагдаж буй шинж тэмдэг нь нэг ижил цаг хугацаанд тохиолдсон/илэрсэн хэсэг бүлэг хүмүүс. Жишээ нь, 1980 оны төрөлтийн үелрэл гэвэл 1980 онд төрсөн бүх хүмүүсийг хамруулан ойлгоно. Мөн 2010 оны их дээд сургууль төгсөгчид нь 2010 оны төгсөлтийн үелрэл гэж ойлгогдоно.

Анх Vincent (1995)-ийн боловсруулан гаргасан “extinct generation” буюу “өнгөрсөн үе” хэмээх арга нь дээрхтэй ижилхэн санаан дээр үндэслэдэг. Энэ нь тухайн үелрэлийн x наснаас хойшхи үелрэл дуусах хүртэлх хугацааны өсөн нэмэгдсэн нас баралтын тооноос x настай үелрэлийн хэмжээг ухрааж тооцдог. Томъёогоор илэрхийлбэл, t хугацаан дахь x настай хүн амын тоо нь:

$$\hat{N}(x, t) = \int_0^{\infty} D^*(x + a, t + a) da \quad (5)$$

байх ба энд, $D^*(x + a, t + a)$ нь $t + a$ үед $x + a$ насандаа нас барсан хүмүүсийн бүртгэгдсэн тоо юм. (Хэрэв тухайн үелрэлийн хувьд шилжих хөдөлгөөн нээлттэй байсан гэж үзвэл, $t + a$ үеийн $x + a$ настай цэвэр шилжигчдийн тоо болох $I^*(x + a, t + a)$ нь $D^*(x + a, t + a)$ -аас хасагдах ёстой)

“Extinct generation” аргын дутагдалтай тал нь амьдралынх нь үе дуусаагүй, одоо амьдарч байгаа үеийнхэний хувьд хэрэглэгдэх боломжгүйд оршиж байна. Гэхдээ үелрэлийн нас баралтын оронд тухайн 1 үеийн нас баралтын тооноос хүн амын тоо хэмжээг тооцож олоход “extinct generation”-аргын логикийг хэрэглэх боломжтой юм. Үүнд дараахь томъёог ашиглана:

$$\hat{N}(y, t) = \int_y^{\infty} D^*(x, t) e^{\int_y^x r^*(a, t) da} dx \quad (6)$$

Энд, $r^*(a, t)$ нь нэг үеийн нас баралтын үзүүлэлтийг t хугацаан дахь y настай үелрэлд тохируулан “засварлагч”. Хэрэв, бүх өсөлтийн хувь тэгтэй тэнцүү бол эндээс тухайн хүн амыг стационарь гэж үзэж болох ба ийм хүн амын хувьд бид ямар нэг насан дахь нас баралтын түвшнийг цаг хугацаанаас хамаарахгүйгээр тогтмол байна гэж авч үзэж болно.

3.3 Загвар ашиглан мэдээллийг үнэлэх ба засварлах

3.3.1 Нас баралтын бүртгэлийн хамрагдалтыг үнэлэх Брассын арга

Мэдээллийн хамрагдалтын түвшинд үнэлгээ хийхэд хүн амын загвар ашиглах анхны оролдлогыг Брасс (1975) хийсэн байдаг. Уг арга нь нас баралтын бүртгэлийн

хамрагдалтын түвшинг үнэлэх зорилготой байжээ. Энэ нь ямар нэг хаалттай хүн ам⁴-ын хувьд тохирох дараахь энгийн тэнцэтгэл дээр үндэслэсэн байсан юм:

$$r = b - d$$

Энд, r , b , d нь тодорхой авч үзэж буй хугацааны өсөлт, төрөлт, нас баралтын бодит коэффициентүүд. Энэ тэнцэтгэл нь мөн x -ээс дээш настай хэсэг хүн амын хувьд ч биелэгдэх боломжтой ба үүнийг томъёолбол:

$$\begin{aligned} r(x+) &= b(x+) - d(x+) \quad \text{эсвэл} \\ b(x+) &= r(x+) + d(x+) \quad (7) \end{aligned}$$

Энд, $r(x+)$, $d(x+)$ нь x ба түүнээс дээш насны хүн амын өсөлтийн коэффициент болон нас баралтын коэффициент, харин $b(x+)$ нь мөн насны хүн амын “төрөлтийн коэффициент”, өөрөөр хэлбэл x настай хүмүүсийн тоог x ба түүнээс дээш насны хүмүүсийн нийт тоонд харьцуулсан харьцаа. Энэхүү “төрөлтийн коэффициент”-ийг дараахь томъёог ашиглан тооцно:

$$b^*(x, t) = \frac{N(x, t)}{N(x+, t)} = \frac{1}{10} \frac{{}_5N_{x-5}(t) + {}_5N_x(t)}{\sum_{a=x, 5}^{\max} {}_5N_a(t)} \quad (8)$$

Брасс энд, хүн амаа стабиль гэж төсөөлсөн, өөрөөр хэлбэл $r(x+) = r$ буюу насны бүлгээр тооцсон өсөлтийн коэффициент нь бүх насанд тогтмол гэсэн төсөөллийг хэрэглэсэн байна. Мөн нас баралтын бүртгэлийн хамрагдалтын түвшин болох R (бүртгэгдсэн тоог бодит тоонд нь харьцуулсан харьцаа) нь бүх насанд тогтмол гэсэн хоёр дахь төсөөллийг хэрэглэжээ. Эндээс, x ба түүнээс дээш насныхны нас баралтын “бодит” коэффициент нь дараахь хэлбэртэй болно:

$$d(x+) = \frac{d^*(x+)}{R}$$

Энд, $d^*(x+)$ нь x ба түүнээс дээш насныхны бүртгэгдсэн нас баралтын коэффициент. Дээрх төсөөллүүдийг (7) тэнцэтгэлд тусгаж, орлуулалт хийвэл дараахь хэлбэртэй болно:

4 Гадаад шилжих хөдөлгөөн бага буюу байхгүй хүн ам

$$b^*(x+) = r + \frac{1}{R} \cdot [d^*(x+)] \quad (9)$$

Хэрэв төсөөллүүд үнэхээр зөв бол энэ тэнцэтгэл бүх насанд биелэх ёстой бөгөөд $b^*(x+)$ ба бүртгэгдсэн нас баралтын коэффициент болох $d^*(x+)$ -ийн хамаарлын шугам шулуун зурагдах ёстой. Брасс нас баралтын бүртгэлийн хамрагдалтын түвшинг үнэлэхдээ $d^*(x+)$ дээрх $b^*(x+)$ -ийн шугаман регресс байгуулахыг санал болгожээ. Энэ тохиолдолд регрессийн шулууны огтлох цэг нь r буюу өсөлтийн коэффициент, налалт нь $1/R$ буюу бүртгэлийн хамрагдалтын түвшин R -ийн урвуу хэмжигдэхүүн байна.

Бодит байдалд, мэдээж (9) тэнцэтгэлээр тодорхойлогдсон цэгүүд нь яг шулуун шугам үүсгэх нь ховор. Нас баралт юмуу хүн амын мэдээлэл дэх насаа буруу мэдээлсэн алдаа нь, ялангуяа ахимаг насныханы хувьд, дээрх цэгүүдийн байршилд илүүтэй нөлөөлж болох юм. Ийм цэгүүд параметрийн тооцоонд хүчтэй нөлөө үзүүлдэг. Иймээс Брасс сүүлийн насны бүлгийнхэний шулуун шугамаас хэт алслагдсан цэгүүдийг үл ойшоохыг зөвлөмж болгосон.

Энэ аргын үр дүн нь стабиль нөхцөл байдал өөрчлөгдөх тохиолдолд, ялангуяа стабиль бус хүн ам буюу үзэгдлийн өөрчлөлт нь түргэн, хурдацтай байх үед ихээхэн мэдрэмтгий байх боломжтой (Martin, 1980). Нас баралт нь огцом буурч байгаа тохиолдлыг авч үзвэл, $r(x+)$ функцын утга

тогтмол байхгүй бөгөөд нас ахих тусам ерөнхийдөө өснө. Иймд, (9) тэнцэтгэлээс тооцоолсон налалт нь биасын нөлөөгөөр нэмэгдэнэ гэсэн үг. Энэ нь бүртгэлтийн хамрагдалтын түвшин R -ийг дутуу үнэлэхэд хүргэж байна. Иймээс, R -ийн тооцоолсон утга нь нас баралт буурч байгаа болон бусад төсөөллүүд үнэн байх нөхцөлд хамрагдалтын бодит түвшин харьцангуй бага хэмжээтэй байна гэж тайлбарлагдана. Шигтгээ 1-ээс Сальвадорын мэдээлэлд Брассын техникийг хэрэглэсэн жишээг харна уу.

3.3.2 Тооллого хооронд шилжин амьдрагсдын мэдээллээс нас баралтын бүртгэлийн хамрагдалтыг үнэлэх аргууд

Престон, Хилл (1980) нар стабиль байх төсөөлөл шаардахгүйгээр нас баралтын бүртгэлийн хамрагдалтыг үнэлэх сонгодог аргыг гаргаж тавьжээ. Энэ нь (2) тэнцэтгэлээр илэрхийлсэн хүн ам зүйн харилцан уялдаанд үндэслэдэг. Уг илэрхийллийг хаалттай хүн амын хувьд авч үзвэл, хоёр дахь тооллогын үелрэлийн хэмжээ нь нэг дэх тооллогын үеийн тухайн үелрэлийн хэмжээ болон тооллого хооронд тухайн үелрэлээс нас барсан хүмүүсийн тоотой дараахь байдлаар холбогддог байна:

$$N_c(2) = N_c(1) - D_c \quad (10)$$

Энд, $N_c(1)$ ба $N_c(2)$ нь үелрэлийн 1 дэх ба 2 дахь тооллогын үеийн бодит тоонууд, харин D_c нь тооллого хооронд тухайн үелрэлээс нас барсан хүмүүсийн бодит тоо.

$$b^*(x+)$$

Шигтгээ 1: Нас баралтын бүртгэлийн хамрагдалтыг үнэлэх Брассын арга

${}_5N_x$ = x -ээс $x+5$ насны, бүртгэгдсэн жилийн дундаж эмэгтэй хүн амын тоо

${}_5D_x$ = судалгааны хугацаанд x -ээс $x+5$ насны хооронд нас барсан хүмүүсийн бүртгэгдсэн тоо

$$N(x+) = \sum_{a=x,5}^{\max} {}_5N_a ; N(x) = \frac{{}_5N_{x-5} + {}_5N_x}{10} ; D(x+) = \sum_{a=x,5}^{\max} {}_5D_a$$

$$d^*(x) = \frac{D(x+)}{N(x+)} ; b^*(x) = \frac{N(x)}{N(x+)}$$

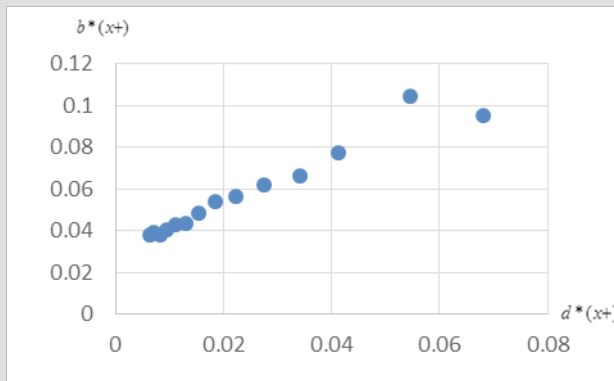
Жишээ: Сальвадор, эмэгтэйчүүд, 2001

x	${}_5N_x$	${}_5D_x$	$N(x+)$	$N(x)$	$D(x+)$	$d^*(x+)$	$b^*(x+)$
0	214,089	6,909					
5	190,234	610	1,060,164	40,432	6,743	0.0064	0.0381
10	149,538	214	869,930	33,977	6,133	0.0070	0.0391
15	125,040	266	720,392	27,458	5,919	0.0082	0.0381
20	113,490	291	595,352	23,853	5,653	0.0095	0.0401
25	91,663	271	481,862	20,515	5,362	0.0111	0.0426
30	77,711	315	390,199	16,937	5,091	0.0130	0.0434
35	72,936	349	312,488	15,065	4,776	0.0153	0.0482
40	56,942	338	239,552	12,988	4,427	0.0185	0.0542
45	46,205	357	182,610	10,315	4,089	0.0224	0.0565
50	38,616	385	136,405	8,482	3,732	0.0274	0.0622
55	26,154	387	97,789	6,477	3,347	0.0342	0.0662
60	29,273	647	71,635	5,543	2,960	0.0413	0.0774
65	14,964	449	42,362	4,424	2,313	0.0546	0.1044
70	11,205	504	27,398	2,617	1,864	0.0680	0.0955
75+	16,193	1,360					

$b^*(x+)$ ба $d^*(x+)$ -ийн хамаарлын дүрслэл:

Регрессийн коэффициентүүд:

$r = 0.0311$ $1/R=1.1002$ буюу $R=0.909$



Эх сурвалж: United Nations, 1993: 135.

Судлаачид нас баралтын бүртгэл ба хоёр тооллогын тус тусын хамрагдалт нь мэдээллийн эх үүсвэрүүдэд ялгаатай байж болно, гэхдээ эх үүсвэр тус бүрийнхээ хүрээнд насны хувьд тогтмол байна гэсэн

төсөөллийг хэрэглэжээ. Ийм тохиолдолд, 1 ба 2 дахь тооллогын үеийн үелрэлийн бүртгэгдсэн тоонууд болох $N_c(1)^*$ ба $N_c(2)^*$ -ууд нь дараахь харилцан хамаарлаар бодит утгуудтайгаа холбогдож болно:

$$N_c(1)^* = N_c(1) \cdot E(1),$$

ба

$$N_c(2)^* = N_c(2) \cdot E(2)$$

Энд, (E1), (E2) нь харгалзан 1 ба 2 дахь тооллогын бүртгэлийн хамрагдалтын түвшин. Мэдээлэгдсэн тооллого хоорондын нас баралт болох D_c^* -г дээрхтэй ижил тэнцэтгэлээр илэрхийлж болно:

$$D_c^* = D_c \cdot R,$$

Энд, R нь тооллого хоорондын нас баралтын бүртгэлийн хамрагдалтын түвшин.

Дээрх үзүүлэлтүүдийг (10) тэнцэтгэлд орлуулж, хувиргавал дараахь байдалтай болно:

$$\frac{N_c^*(1)}{E(1)} = \frac{N_c^*(2)}{E(2)} + \frac{D_c^*}{R}$$

буюу,

$$\frac{N_c^*(1)}{N_c^*(2)} = \frac{E(1)}{E(2)} + \frac{E(1)}{R} \cdot \frac{D_c^*}{N_c^*(2)} \quad (11)$$

Брассын аргад дурьдаж байсан шиг, $E(1)/E(2)$ ба $E(1)/R$ утгуудыг өөр өөр үелрэлийн мэдээллээс гаргасан энгийн шугаман регресс ашиглах замаар тооцож болно. Хоёр мэдэгдэж буй үзүүлэлт нь $N_c^*(1)/N_c^*(2)$ ба $D_c^*/N_c^*(2)$ байна. $E(1)/E(2)$ гэсэн огтлох цэг нь хоёр тооллогын хамрагдалтын харьцаа, $E(1)/R$ буюу налалт нь нэг дэх тооллогын хамрагдалтын түвшинг нас баралтын бүртгэлийн хамрагдалтын түвшинд харьцуулсан харьцаа. Энэ тооцоонд хэрэглэгдсэн үелрэлүүд нь заагдсан хугацааг хамарсан, эсвэл нээлттэй төгсгөлтэй, өөрөөр хэлбэл тодорхой сонгосон жил болон түүнээс арай өмнөх жилд төрсөн байж болно.

Энд нэг зүйлийг тэмдэглэхэд, уг арга нь яг абсолют биш боловч, нэг дэх тооллогод харьцуулсан байдлаар бүртгэлийн хамрагдалтын түвшинг үнэлдэг. Энэ харьцаа нь нас баралтыг тооцоход ямар нэг асуудал болохгүй, учир нь нас баралтын түвшинг засварлахын тулд зөвхөн холбогдох алдааг мэдэхэд л хангалттай. Харин нас баралтын

бүртгэлийн хамрагдалтын үнэлгээг гажуудуулахад хүргэдэг нас хэтрүүлж хэлэх нөхцөл байдалд ихээхэн мэдрэмжтэй байдаг нь анхаарууштай асуудал юм. Нас тодорхойлох асуудал нь ерөнхийдөө харьцангуй бэрхшээлтэй байдаг Латин Америкийг бодвол Зүүн Азийн хүн амын хувьд энэ арга илүү тохиромжтой гэж үздэг.

Hill (1987) энэ аргын нас буруу мэдээлэх алдаанаас хэт хамааралтай байдлыг арай багасган сайжруулсан. Өөрөөр хэлбэл, тооллогоос тооллогын хоорондох үелрэлийн хэмжээнд гарсан өөрчлөлтийн оронд насны бүлгийн хэмжээнд гарсан өөрчлөлтийг анхааран судалсан. Энэ нь тооллогын хамрагдалтын хүрээ ба нас баралтын бүртгэлийн дутуу хамрагдалтын түвшингүүдийн зөрүүг ашиглан (7) тэнцэтгэлд үндэслэдэг.

3.3.3 Насжилтын загвар хүснэгт ашиглан ахимаг насныханы нас баралтыг үнэлэх

Бидний урьд нь авч үзэж байснаар, насаа буруу мэдээлэх явдал нь ахимаг насныханы нас баралтын коэффициентийг тооцоолоход гардаг үндсэн бэрхшээлтэй асуудлуудын нэг юм. Эдгээр наснууд нь нас баралт судлалд улам чухал болсоор байна, учир нь дэлхийн хүн амын нэлээд өндөр хувь нь уг насанд харъяалагддаг билээ. Азаар, насаа буруу мэдээлэх явдал нь нийт хүн амын нас баралтын ерөнхий коэффициентэд нөлөөлдөггүй байна. Хэрэв, насаа буруу мэдээлэх явдал нь зөвхөн ахимаг насныханд, тодруулбал 65 ба түүнээс дээш насныханд тохиолддог гэвэл, 65 ба түүнээс дээш насны хүн амын нас баралтын коэффициент дээрх нөлөөнд орохгүй. Ахимаг насныханы нас баралтыг үнэлэхдээ бүртгэгдсэн, алдаатай коэффициентуудыг насжилтын ямар нэг загвар хүснэгт ашиглан тооцсон нас баралтын коэффициентуудаар орлуулж болох юм (Elo, Preston, 1994).

Энэ арга нь дараахь төсөөллүүдэд үндэслэнэ. Нэгдүгээрт, тодорхой насан дахь мэдээлэгдсэн нас баралтын коэффициент нь бодит коэффициенттойгоо адил гэж үзнэ. Хоёр дахь төсөөлөл нь, насаа буруу мэдээлэх явдал тооллогуудын үед гардаг гэж үзвэл, хоёр тооллогын үед алдааны хандлага

пропорционалиар адилхан байна гэж авах ба эндээс, тооллого хоорондын насны бүлгээр тооцсон өсөлтийн коэффициентууд үнэн зөв хэвээр үлдэнэ.

Уг аргын эхний шатанд хийгдэх үйлдэл бол тухайн нас хүртэлх бүх насны мэдээлэл нийцтэй байх тийм Y гэсэн хамгийн өндөр насыг тодорхойлох явдал юм. Үүнд, жишээлбэл, энэ бүлгийн эхэнд өгүүлж байсан нийцэмжтэй байдлыг шалгах аргуудын нэгийг ашиглаж болно. Уг наснаас доошхи мэдээлэгдсэн бүх насны бүлгээрх нас баралтын коэффициентүүд нь эцсийн бидний гаргаж авахыг хүсч буй насжилтын хүснэгтийг зохиоход ашиглагдана. Хоёр дахь шатанд ахимаг насныханы хувьд насжилтын загвар хүснэгтийн системийг ашиглах шаардлагатай. Нэлээд хэдэн загвар хүснэгтийн систем өгөгдсөн үед, Y наснаас хойшхи мэдээлэгдсэн нас баралтын ерөнхий коэффициент нь нийцэмжтэй байх тийм системийн хүрээнд л нэг хүснэгтийг байгуулах ёстой. Y -ээс хойшхи насныхны нас баралтын коэффициентын илэрхийлэл нь дараахь байдалтай байна:

$$DR_{Y+} = \frac{\int_Y^{\infty} N(a)\mu(a)d}{\int_Y^{\infty} N(a)d} \quad (12)$$

$N(a)$ нь Y -ээс дээш насанд буруу мэдээлэгддэг гэж тооцогдож байгаа учраас, түүнийг дараахь илэрхийллээр (8.1 тэнцэтгэл) орлуулж болох юм:

$$N(a) = N(Y)e^{-\int_Y^a r(x)d} \frac{p(a)}{p(Y)}$$

Эндээс,

$$DR_{Y+} = \frac{\int_Y^{\infty} e^{-\int_Y^a r(x)d} p(a)\mu(a)d}{\int_Y^{\infty} e^{-\int_Y^a r(x)d} p(a)d} = \frac{\int_Y^{\infty} e^{-\int_Y^a r(x)d} d(a)d}{\int_Y^{\infty} e^{-\int_Y^a r(x)d} p(a)d} \quad (13)$$

болох ба $r(x)$ нь x насан дахь мэдээлэгдсэн өсөлтийн коэффициент, $d(a)$ нь a насандахь загвар насжилтын хүснэгтийн нас баралт, $p(a)$ нь загвар насжилтын хүснэгтийн a насанд шилжин амьдрагчдын тоо юм. Тэгвэл илэрхийлэл (13) нь мэдээлэгдсэн насны

бүлгээр тооцсон өсөлтийн коэффициентууд болон Y -ээс дээш насныхны нас баралтын ерөнхий коэффициент нь нийцтэй байх тийм насжилтын загвар хүснэгтийг сонгоход ашиглагдах болно. Насжилтын загвар хүснэгтэд дундаж наслалт нэмэгдэх тутам, тэнцэтгэлийн баруун гар талын утга нь зүүн гар талын ажиглагдсан нас баралтын ерөнхий коэффициенттэй тэнцтэл багасна.

АШИГЛАСАН МАТЕРИАЛУУД:

Brass W. 1975. *Methods for Estimating Faculty and Mortality from Limited and Defective Data*, Chapel Hill, North Carolina: Laboratories for Population Statistics, Carolina Population Centre

Caldwell J. 1966. "A Study of age misstatement among young children in Ghana", *Demography* 3(2): 477-302.

Chandrasekar, C.and W. Edwards Deming, March 1949. "On a Method of Estimating Birth and Death Rates and the Extent of Registration," *Journal of the American Statisticl Association*, Vol. 44: pp. 101-115.

Elo, I.T. and S.H. Preston. 1994. "Estimating African-American Mortality from Inaccurate Data." *Demography* 31:427-458. PMID: 7828765

Ewbank, Douglas C. 1981. *Age Misreporting and Age Selective Underenumeration: Sources, Patterns, and Consequences for Demographic Analysis*. Committee on Population and Demography, Report No. 4. Washington, DC: National Academy Press.

Newcombe, H. B. *Handbook of Record Linkage*, OUP: New York, 1988.

Preston S. and Hill K. 1980. "Estimating the completeness of Death Registration", *Population Studies*, Vol.34(2), pp.349-366

Preston, Samuel H., Irma T. Elo, Ira Rosenwaike, and Mark Hill. 1996. "African-American mortality at older ages: Results of a matching study", *Demography* 33(2): 193-209.