

# Байгалийн шинжлэх ухаан, математик, статистикийн мэргэжлээр төгсөгчдийн тооны хандлага

Г.Махгал

2025 оны 10-р сарын 14

Энэхүү ажлаар Монгол улсын их, дээд сургуулийг байгалийн ухаан, математик, статистикийн мэргэжлээр төгсөгчдийг тоог Үндэсний статистикийн хорооны албан эх сурвалж дээрх өгөгдөлд тулгуурлан авч үзэх бөгөөд улмаар төгсөгчдийн тооны чиг хандлага ба ойрын жилүүдийн прогноз гаргана.

## Агуулга

Шаардлагатай багцууд	1
1 Өгөгдөл	2
2 Төгсөгчдийн тоо	4
3 Төгсөгчдийн тооны хандлага	5
4 Төгсөгчдийн тооны загварчлал	6
5 Төгсөгчдийн тооны прогноз	7
Дүгнэлт	8
Ашигласан материал	9

## Шаардлагатай багцууд

Шаардлагатай багцуудыг дараах байдлаар урьдчилан суулгана.

```
pip install NS01212v1 tabulate matplotlib scipy statsmodels
```

# 1 Өгөгдөл

Үндэсний статистикийн хороо (ҮСХ) нь 1212.mn [1] веб сайтаараа дамжуулан Монгол улсын нийгэм, эдийн засгийн үзүүлэлтийг илэрхийлэх тоон мэдээллийг нээлттэй тархаадаг. Тус веб сайт дээр Монгол улсын их, дээд сургуулийг төгсөгчдийн тоог мэргэжлийн салбар бүрээр, жил жилээр нэгтгэн мэдээлжээ. Бид үүнээс байгалийн ухаан, математик, статистикийн мэргэжлээр төгсөгчдийн тоог авч үзнэ.

Байгалийн шинжлэх ухаан, математик, статистикийн мэргэжлээр төгсөгчдийн тоог ҮСХ-ны 1212.mn [2] веб сайтаас татаж авна.

ҮСХ-ны 1212.mn дээрээс өгөгдөл татахад Python хэлний NSO1212v1 багц [3] ашиглана. Багцыг командын мөрөөс `pip install NSO1212v1` тушаалаар суулгана. Тус багц нь ҮСХ-ны 1212 нээлттэй өгөгдлийн API-ийн хувилбар 1-ийг дэмжинэ.

```
import NSO1212v1 as nso
```

ҮСХ-ны нээлттэй өгөгдлийн APIv1 нь өгөгдөл татахад зөвхөн хүснэгтийн ID төдийгүй сектор болон дэд секторын ID-ийг ч шаарддаг.

```
nso.sectors()[["id", "text"]].rename(columns={"id": "ID", "text": "Сектор"})
```

Хүснэгт 1: Үндсэн секторууд, тэдгээрийн ID

	ID	Сектор
0	Education, health	Боловсрол, эрүүл мэнд
1	Regional development	Бүсчилсэн хөгжлийн үзүүлэлтүүд
2	Society, development	Нийгэм, хөгжил
3	Historical data	Түүхэн Статистик
4	Industry, service	Үйлдвэрлэл, үйлчилгээ
5	Labour, business	Хөдөлмөр, бизнес
6	Population, household	Хүн ам, өрх
7	Economy, environment	Эдийн засаг, байгаль орчин

Боловсрол, эрүүл мэндийн салбар “Education, health” id-тай байна.

```
nso.subsectors(sector_id="Education, health")[["id", "text"]].rename(columns={"id": "ID", "text": "Дэд сектор"})
```

Хүснэгт 2: Дэд секторууд, тэдгээрийн ID

	ID	Дэд сектор
0	General indicators for Education	Боловсролын ерөнхий үзүүлэлт
1	General educational schools	Ерөнхий боловсролын сургууль
2	Universities, institutes and colleges	Их дээд сургууль
3	Vocational education	Мэргэжлийн боловсрол
4	Disease	Өвчлөл
5	Pre-school education	Сургуулийн өмнөх боловсрол
6	Births, deaths	Төрөлт, нас баралт
7	Health insurance	Эрүүл мэндийн даатгал
8	Main indicators for Health sector	Эрүүл мэндийн үндсэн үзүүлэлт

Их дээд сургууль гэсэн дэд салбарын id нь “Universities, institutes and colleges” байна.

Одоо дээрх дэд салбарт хамаарах хүснэгтүүдийн жагсаалтыг авна.

```
nso.tables(
  sector_id="Education, health", subsector_id="Universities, institutes and colleges"
)[["id", "text"]].rename(columns={"id": "ID", "text": "Хүснэгт"})
```

Хүснэгт 3: Хүснэгтүүд, тэдгээрийн ID

ID	Хүснэгт
0 DT_NSO_2001_013V1.px	ИХ, ДЭЭД СУРГУУЛЬ, КОЛЛЕЖИД СУРАЛЦАГЧИД, хүйс,...
1 DT_NSO_2001_013V2.px	ИХ, ДЭЭД СУРГУУЛЬ, КОЛЛЕЖИД СУРАЛЦАГЧИД, мэргэ...
2 DT_NSO_2001_014V1.px	ИХ, ДЭЭД СУРГУУЛЬ, КОЛЛЕЖИЙГ ТӨГСӨГЧИД, мэргэж...
3 DT_NSO_2001_015V2.px	ДЭЭД БОЛОВСРОЛЫН СУРГАЛТЫН БАЙГУУЛЛАГАД СУРАЛЦ...
4 DT_NSO_2001_016V1.px	ИХ, ДЭЭД СУРГУУЛЬ, КОЛЛЕЖИЙН ҮНДСЭН БАГШИЙН ТО...
5 DT_NSO_2001_024V1.px	ИХ, ДЭЭД СУРГУУЛЬ, КОЛЛЕЖИЙГ ТӨГСӨГЧИД, өмчийн...
6 DT_NSO_2001_024V2.px	ДОТООДЫН ИХ, ДЭЭД СУРГУУЛЬ, КОЛЛЕЖИЙГ ТӨГСӨГЧИ...
7 DT_NSO_2001_024V3.px	ДОТООДЫН ИХ, ДЭЭД СУРГУУЛЬ, КОЛЛЕЖИД СУРАЛЦАГЧ...

Жагсаалтаас их, дээд сургууль, коллежийг төгсөгчид, мэргэжлийн чиглэлээр гэсэн хүснэгтийн id дугаар нь “DT\_NSO\_2001\_014V1.px” гэж харагдана.

Одоо хүснэгтийн мета мэдээллийг татаж авна.

```
table_meta_data = nso.table(
  sector_id="Education, health",
  subsector_id="Universities, institutes and colleges",
  table_id="DT_NSO_2001_014V1.px"
)
```

Мета мэдээллээс мэргэжлийн чиглэл хувьсагчийн утгууд, тэдгээрийн кодыг дараах байдлаар олж харна.

```
table_meta_data["variables"][0]["codes"].rename(columns={"values": "Код", "valueTexts": "Мэргэжлийн чиглэл"})
```

Хүснэгт 4: Мэргэжлийн чиглэл хувьсагчийн утгууд, тэдгээрийн код

Код	Мэргэжлийн чиглэл
0 0	Бүгд
1 1	Боловсрол
2 2	Урлаг, хүмүүнлэг
3 3	Нийгмийн шинжлэх ухаан, мэдээлэл, сэтгүүл зүй
4 4	Бизнес, удирдахуй, хууль, эрх зүй
5 5	Байгалийн шинжлэх ухаан, математик, статистик
6 6	Мэдээлэл, харилцаа холбооны технологи
7 7	Инженер, үйлдвэрлэл, барилга угсралт
8 8	Хөдөө аж ахуй, ой, загасны аж ахуй, мал эмнэлэг
9 9	Эрүүл мэнд, нийгмийн хамгаалал
10 10	Үйлчилгээ

“Байгалийн шинжлэх ухаан, математик, статистик” мэргэжлийн чиглэлийн код “5” байна.

Ийнхүү өгөгдөл татаж авахад шаардлагатай бүх мэдээлэл тодорхой боллоо. Одоо өгөгдөл татаж авна.

```
data = nso.data(
    sector_id="Education, health",
    subsector_id="Universities, institutes and colleges",
    table_id="DT_NSO_2001_014V1.px",
    query={
        "query": [
            {
                "code": "Мэргэжлийн чиглэл",
                "selection": {
                    "filter": "item",
                    "values": ["5"]
                }
            }
        ],
        "response": {
            "format": "json-stat2"
        }
    }
)
```

Тус функц серверээс ирсэн объект доторх өгөгдлийг эцсийн хэрэглэгчдэд таатай байдлаар эмхэлж өгдөг. Тэр нь df нэртэй, pandas.DataFrame бүтэцтэй элементэд агуулагдана. Ингээд data.df датафреймыг “raw\_data.csv” файлд хадгална.

```
data["df"].to_csv("raw_data.csv", index=False)
```

Одоо өгөгдлөөс зөвхөн “Он” ба “Хувьсагч” хоёр баганыг ялгаж аваад бусдыг нь гээгнэ. Бас хувьсагчдын нэрсийг year, graduates болгон солино.

```
import pandas as pd

data = data["df"][["Он", "Хувьсагч"]]
data.rename(columns={"Он": "year", "Хувьсагч": "graduates"}, inplace=True)
data["year"] = pd.to_numeric(data["year"], errors='coerce')
```

Ийнхүү ҮСХ-ноос мэдээлсэн байгалийн шинжлэх ухаан, математик, статистик мэргэжлээр жил бүр төгсөгчдийн тоог агуулсан өгөгдөлтэй боллоо.

## 2 Төгсөгчдийн тоо

ҮСХ-ны нээлттэй өгөгдлийн сервисээс татан авсан өгөгдлийг Хүснэгт 5 дээр харууллаа.

data

Хүснэгт 5: Байгалийн шинжлэх ухаан, математик, статистик мэргэжлээр жил бүр төгсөгчдийн тоо

	year	graduates
0	2024	424
1	2023	522

Хүснэгт 5: Байгалийн шинжлэх ухаан, математик, статистик мэргэжлээр жил бүр төгсөгчдийн тоо

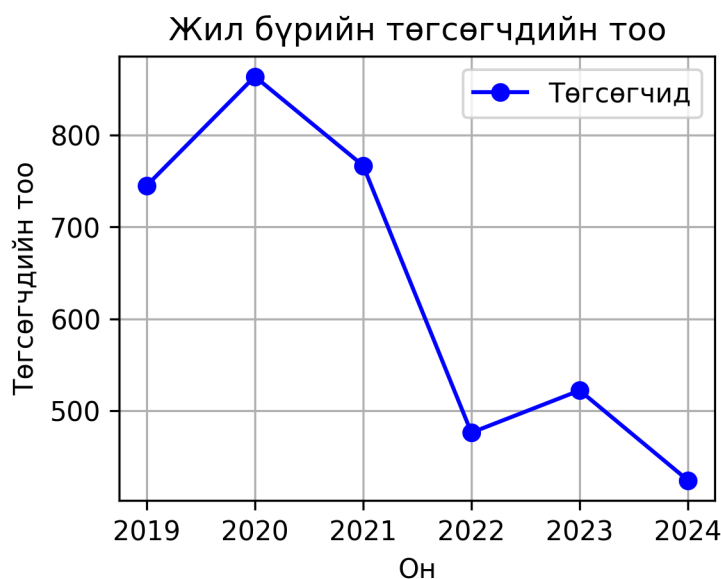
	year	graduates
2	2022	476
3	2021	767
4	2020	864
5	2019	745

Мөн тус өгөгдлийг шугаман диаграммаар дүрслэв. Үүнийг Зураг 1 дээрээс харж болно.

```
import matplotlib.pyplot as plt

plt.figure(figsize=(4,3))
plt.plot(data["year"], data["graduates"], marker='o', linestyle='-', color='b', label='Төгсөгчид')
plt.xlabel("Он")
plt.ylabel("Төгсөгчдийн тоо")
plt.title("Жил бүрийн төгсөгчдийн тоо")
plt.grid(True)
plt.legend()
plt.show()
```

Зураг 1: Байгалийн шинжлэх ухаан, математик, статистикийн мэргэжлээр төгсөгчдийн тоо



Шугаман диаграммаас төгсөгчдийн тоо буурах хандлагатай гэсэн таамаглал гарна.

### 3 Төгсөгчдийн тооны хандлага

Өмнөх хэсэгт байгалийн шинжлэх ухаан, математик, статистикийн мэргэжлээр төгсөгчдийн тоо буурах хандлагатай гэсэн таамаглал дэвшүүлсэн. Одоо түүнийг шалгана. Статистик таамаглал шалгалтын арга зүйн үүднээс хандлага тогтмол гэсэн тэг таамаглалын эсрэг хандлага буурна гэсэн өрсөлдөгч таамаглал дэвшүүлнэ. Улмаар тэг таамаглалыг няцаавал өрсөлдөгч таамаглалыг баталсан явдал болно.

Ийм таамаглалыг корреляц бас регресс гэх мэтчилэн янз бүрийн аргаар шалгаж болохоос корреляцийн коэффициентын тусламжтай шалгах аргыг сонголоо. Энэ тохиолдолд таамаглалыг дараах хэлбэрээр томъёолно.

$$H_0 : \text{cor}(\text{graduates}, \text{year}) = 0$$
$$H_1 : \text{cor}(\text{graduates}, \text{year}) < 0$$

Мөн түүврийн хэмжээ бага бас тархалт нормал биш тул параметрийн бус корреляц [4] ашиглана.

```
from scipy import stats

rho, p_one_sided = stats.spearmanr(
    data["year"], data["graduates"],
    alternative="less"           # корреляц тэгээс бага
)
```

Корреляц -0.771 гэж гарав. Улмаар  $p\text{-value} = 0.036 < \alpha = 0.05$  гэж гарсан нь хандлага буурч байна гэсэн өрсөлдөгч таамаглалыг баталж буй явдал юм.

```
conclusion.append("Байгалийн шинжлэх ухаан, математик, статистикийн мэргэжлээр төгсөгчдийн тоо буурч байна.")
```

## 4 Төгсөгчдийн тооны загварчлал

Төгсөгчдийн тооны прогноз гаргахын тулд үүнийг оноос хамааруулан загварчлах шаардлагатай. Прогноз гаргахад хугацаан цувааны загвар эсвэл регрессийн загвар ашигладаг. Ийнхүү төгсөгчдийн тоог жилээс хамааруулсан регрессийн загвар авч үзнэ.

Загвар дахь хамааран хувьсагч нь тоо ширхэг илэрхийлсэн. Энэ байдалд тохирох загвар бол Пуассоны регрессийн загвар [5] юм.

```
import statsmodels.api as sm
import statsmodels.formula.api as smf

model_poisson = smf.glm(
    formula="graduates ~ year",
    data=data,
    family=sm.families.Poisson()
).fit()
```

Гэхдээ Пуассоны регресс overdispersion [5] буюу хэт дисперстэй өгөгдөлд тохирохгүй. Өгөгдөл хэт дисперстэй эсэхийг шалгадаг R програмын AER багц дахь `dispersiontest()` [6] шиг бэлэн функц байхгүй тул үүнийг “гар аргаар” ажиллуулна.

```
deviance = model_poisson.deviance
df_resid = model_poisson.df_resid
dispersion = deviance / df_resid
```

Дисперсийн шинжүүрийн утга 17.166 гэж 1.5 гэсэн босгоос их гарсан тул өгөгдлийг хэт дисперстэй гэж дүгнэнэ.

```
conclusion.append("Байгалийн шинжлэх ухаан, математик, статистикийн мэргэжлээр төгсөгчдийн тоонд " +
    "хэт дисперстэй тархалтын загвар тохирно.")
```

Хэт дисперстэй хувьсагчийг загварчлах нэг тархалт бол урвасан бином тархалт бөгөөд ийнхүү урвасан бином загвар [5] ашиглана.

```
model_nb = smf.glm(  
    formula="graduates ~ year",  
    data=data,  
    family=sm.families.NegativeBinomial(alpha=1.0) # alpha: dispersion parameter  
) .fit()
```

Прогноз гаргахад шаардлагатай регрессийн загвар бэлэн болсон тул одоо ойрын хэдэн жилийн төгсөгчдийн тооны прогноз гаргах руу чиглэнэ.

## 5 Төгсөгчдийн тооны прогноз

ҮСХ-ны мэдээ 2024 оноор дуусгавар болсон байсан. Иймээс 2025 оноос эхлэн гурван жилийн прогноз гаргана.

```
# Ирээдүйн жилүүд  
future_years = pd.DataFrame({  
    "year": [2025, 2026, 2027]  
})  
  
# Ирээдүйн төгсөгчдийн тоог прогнозлох  
future_years["predicted"] = model_nb.predict(future_years)  
  
round(future_years)
```

Хүснэгт 6: Төгсөгчдийн тооны ойрын ирээдүйн прогноз

	year	predicted
0	2025	378.0
1	2026	329.0
2	2027	286.0

```
# Ирээдүйн төгсөгчдийн тоо  
pred_2027 = round(future_years.loc[future_years["year"] == 2027, "predicted"].values[0])  
  
conclusion.append("Байгалийн шинжлэх ухаан, математик, статистикийн мэргэжлээр төгсөгчдийн тоо " +  
    f"2027 онд {pred_2027} байна гэсэн прогноз гарсан.")
```

```
# Бүх жил (2019–2027) дэх прогноз  
all_years = pd.DataFrame({"year": list(range(2019, 2028))})  
all_years["predicted"] = model_nb.predict(all_years)  
  
# Диаграмм  
plt.figure(figsize=(4,3))  
  
# Ажиглагдсан утгууд  
plt.scatter(data["year"], data["graduates"], color="blue", label="Ажиглалт", zorder=3)  
  
# Загвараар гарсан үнэлгээ (2019–2027)  
plt.plot(  
    all_years["year"], all_years["predicted"],  
    color="darkgreen", linewidth=2, label="Загвар (2019–2027)"
```

```

)

# Ирээдүйн прогноз
future_preds = model_nb.predict(future_years)
plt.scatter(future_years["year"], future_preds, color="red", label="Прогноз (2025–2027)", zorder=4)

# Диаграммын тэнхлэгүүд
plt.xlabel("Он")
plt.ylabel("Төгсөгчид")
plt.xlim(2018, 2028)
plt.ylim(min(data["graduates"].min(), all_years["predicted"].min()) * 0.8,
         max(data["graduates"].max(), all_years["predicted"].max()) * 1.2)

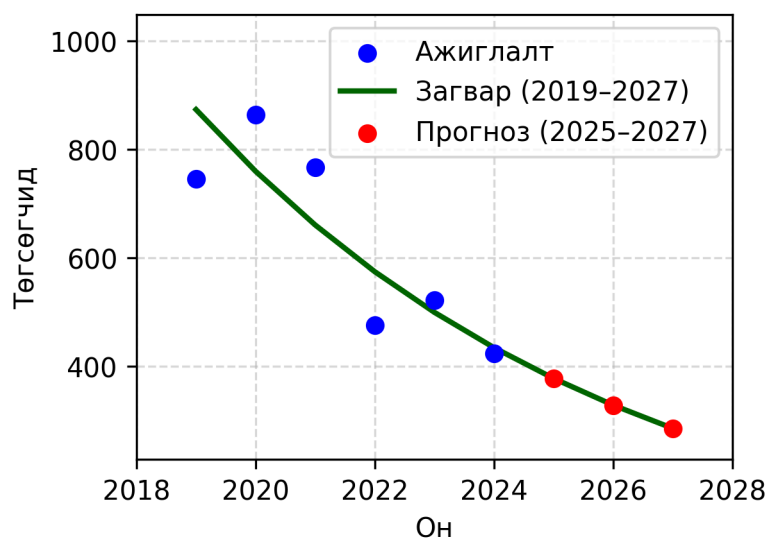
# Legend
plt.legend(loc="upper right")

# Grid
plt.grid(True, linestyle="--", alpha=0.5)

plt.show()

```

Зураг 2: Төгсөгчдийн тоо ба урвасан бином регресс



## Дүгнэлт

```

from IPython.display import display_latex, Latex

latex_code = "\\begin{enumerate}\\n" + "\\n".join(f"\\item {t}" for t in conclusion) + "\\n\\end{enumerate}"
display_latex(Latex(latex_code))

```

1. Байгалийн шинжлэх ухаан, математик, статистикийн мэргэжлээр төгсөгчдийн тоо буурч байна.
2. Байгалийн шинжлэх ухаан, математик, статистикийн мэргэжлээр төгсөгчдийн тоонд хэт дисперс-тэй тархалтын загвар тохирно.
3. Байгалийн шинжлэх ухаан, математик, статистикийн мэргэжлээр төгсөгчдийн тоо 2027 онд 286 байна гэсэн прогноз гарсан.



## Ашигласан материал

- [1] Үндэсний Статистикийн Хороо, “1212.mn — статистикийн мэдээллийн нэгдсэн сан.” Үндэсний Статистикийн Хороо, 2025. Available: <https://www.1212.mn/>
- [2] Монгол Улсын Үндэсний Статистикийн Хороо, “Боловсрол, эрүүл мэндийн салбарын статистик мэдээлэл.” 2001. Available: [https://www.nso.mn/mn/statcate/table-view/Education,%20health/Universities%2C%20institutes%20and%20colleges/DT\\_NSO\\_2001\\_014V1.px](https://www.nso.mn/mn/statcate/table-view/Education,%20health/Universities%2C%20institutes%20and%20colleges/DT_NSO_2001_014V1.px)
- [3] M. Ganbold, *NSO1212v1: National statistical office of mongolia’s open data API v1 handler for python*. (Oct. 08, 2025). Available: <https://pypi.org/project/NSO1212v1/>
- [4] P. Schober, T. R. Vetter, and G. Haller, “Correlation coefficients: Appropriate use and interpretation,” *Anesthesia & Analgesia*, vol. 126, no. 5, pp. 1763–1768, 2018, doi: [10.1213/ANE.0000000000002864](https://doi.org/10.1213/ANE.0000000000002864).
- [5] P. Roback and J. Legler, *Beyond multiple linear regression: Applied generalized linear models and multilevel models in r. Poisson regression*. CRC Press, 2021. Available: <https://bookdown.org/robback/bookdown-BeyondMLR/ch-MLRreview.html>
- [6] C. Kleiber and A. Zeileis, *Applied econometrics with R*. New York: Springer-Verlag, 2008. doi: [10.1007/978-0-387-77318-6](https://doi.org/10.1007/978-0-387-77318-6).