Понятия и формули:

- ## Централна гранична теорема

  Ако случайни извадки от n наблюдения са взети от популация с различно от нормалното разпределение, но с крайно средно μ и стандартно отклонение σ и ако n е голямо, то извадковото разпределение на извадковото средно $\bar{x}$ е приблизително нормално разпределено със следните средно и стандартно отклонение: $\mu_{\bar{x}} = \mu,\ \sigma_{\bar{x}} = \sigma/\sqrt{n}$. Приближението става по-точно с увеличаването на големината на извадката n. Теоремата може да се запише за сумата от наблюденията: $\sum_{i=1}^{n} x_i \sim \mathrm{N}(n\mu, n\sigma^2),\ n \to \infty$. Извадковото разпределение на извадковото средно е точно нормално разпределено, независимо от големината на извадката, когато извадката е от популация с нормално разпределение.

- Извадково разпределение на извадковото средно $\bar{x}$

  – Ако случайна извадки от n наблюдения е взета от популация със средно μ и стандартно отклонение σ, то извадковото средно $\bar{x}$ ще има следните средно и стандартно отклонение: $\mu_{\bar{x}} = \mu,\ \sigma_{\bar{x}} = \sigma/\sqrt{n}$.

  – Ако популацията има нормално разпределение, то $\bar{x}$ има нормално разпределение, независимо от големината на извадката.

  – Ако популацията няма нормално разпределение, то $\bar{x}$ има приблизително нормално разпределение при големи извадки (по Централна гранична теорема). За повечето популации достатъчно голяма извадка е при $n = 25$.

- Свойства на извадковото разпределение на извадковата пропорция $\hat{p}$

  – Ако случайна извадки от n наблюдения е взета от биномна популация с параметър p, то извадковата пропорция $\hat{p} = x/n$ ще има следните средно и стандартно отклонение: $\mu_{\hat{p}} = p,\ \sigma_{\hat{p}} = \sqrt{pq/n},\ q = 1 - p$.

  – Ако големината на извадката е голяма, то разпределението на извадковата пропорция $\hat{p}$ може да бъде приближено чрез нормално разпределение. Приближението е адекватно, ако $\mu_{\hat{p}} \pm 2\sigma_{\hat{p}}$ попадат в интервала между 0 и 1.

- Свойства на извадково разпределение на разликата $(\bar{x}_1 - \bar{x}_2)$ между две извадкови средни

  Когато независими извадки от n_1 и n_2 наблюдения са избрани от популации със средни μ_1 и μ_2 и дисперсии σ_1^2 и σ_2^2, извадковото разпределение на разликата $(\bar{x}_1 - \bar{x}_2)$ ще има следните свойства:

  – средното и стандартното отклонение на $(\bar{x}_1 - \bar{x}_2)$ ще са: $\mu_{(\bar{x}_1 - \bar{x}_2)} = \mu_1 - \mu_2,\ \sigma_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}$.

  – Ако популациите имат нормално разпределение, то $(\bar{x}_1 - \bar{x}_2)$ има нормално разпределение, независимо от големините на извадките.

  – Ако популациите нямат нормално разпределение, то $(\bar{x}_1 - \bar{x}_2)$ има приблизително нормално разпределение, когато са големи извадките, поради Централна гранична теорема.

- Свойства на извадковото разпределение на разликата $(\hat{p}_1 - \hat{p}_2)$ между две извадкови пропорции

  Когато независими извадки от n_1 и n_2 наблюдения са избрани от биномни популации с параметри p_1 и p_2, то извадковото разпределение на разликата $(\hat{p}_1 - \hat{p}_2) = (x_1/n_1 - x_2/n_2)$ ще има следните свойства:

  – средното и стандартното отклонение на $(\hat{p}_1 - \hat{p}_2)$ ще са: $\mu_{(\hat{p}_1 - \hat{p}_2)} = p_1 - p_2,\ \sigma_{(\hat{p}_1 - \hat{p}_2)} = \sqrt{p_1 q_1/n_1 + p_2 q_2/n_2}$.

  – Ако големините на извадките са големи, то разпределението на разликата $(\hat{p}_1 - \hat{p}_2)$ между двете извадкови пропорции може да бъде приближено чрез нормално разпределение. Приближението е адекватно, ако $(p_1 - p_2) \pm 2\sigma_{(\hat{p}_1 - \hat{p}_2)}$ попадат в интервала между -1 и 1.

Зад.1 Да предположим, че имаме извадка от $n = 25$ наблюдения от популация със средно $\mu = 8$ и стандартно отклонение $\sigma = 0.6$. Намерете приближение на вероятността, че извадковото средно $\bar{x}$ ще бъде:

а) по-малко от 7.9; б) по-голямо от 7.9; в) ще лежи в границите на 0.1 от популационното средно $\mu = 8$

Отговор: а) 0.2033; б) 0.7967; в) 0.5934

Зад.2 Направена е случайна извадка от n наблюдения от популация със средно μ и дисперсия σ^2. Намерете средното и стандартното отклонение на извадъчното средно за:

а) $n = 25, \mu = 10, \sigma^2 = 9$; б) $n = 100, \mu = 5, \sigma^2 = 4$; б) $n = 8, \mu = 120, \sigma^2 = 1$

Отговор: а) 10, 0.6; б) 5, 0.2; в) 120, 0.3536

Зад.3 Имаме условието от предишната задача.

а) Ако популациите са нормални, какво е разпределението на $\bar{x}$ за трите подточки?

б) Ако популациите не са нормални, какво можем да кажем за разпределението на $\bar{x}$ за трите подточки?

Отговор: а) нормално; б) приблизително нормално за а) и б)

Зад.4 Направена е случайна извадка от $n = 5$ наблюдения от нормална популация със средно $\mu = 1$ и стандартно отклонение $\sigma = 0.36$.

а) Намерете средното и стандартното отклонение на извадъчното средно за $\bar{x}$

б) намерете вероятността, че $\bar{x}$ ще надхвърли 1.3

в) намерете вероятността, че $\bar{x}$ ще е по-малко от 0.5

г) намерете вероятността, че $\bar{x}$ ще се отклонява от популационното средно с повече от 0.4.

Отговор: а) 1, 0.161; б) 0.0314; в) ≈ 0; г) 0.0132

Зад.5 Направена е случайна извадка от $n = 25$ наблюдения от нормална популация със средно $\mu = 106$ и стандартно отклонение $\sigma = 12$.

а) Намерете средното и стандартното отклонение на извадъчното средно за $\bar{x}$

б) намерете вероятността, че $\bar{x}$ ще надхвърли 110

в) намерете вероятността, че $\bar{x}$ ще се отклонява от популационното средно с не повече от 4.

Отговор: а) 106, 2.4; б) 0.0475; в) 0.9050;

Зад.6 Wall Street Journal (March 20, 1985) докладва за проучване при 313 деца, на възраст между 14 и 22 години, избрани сред децата на директори на най-големите национални компании. Попитани да определят най-добрата страна на това да бъдеш част от привилегирована група, 55% споменават материалните и финансовите предимства. Опишете извадковото разпределение на извадковата пропорция $\hat{p}$ на децата, поставящи материалните предимства като най-добрата страна на техния привилегирован живот.

Отговор: $\sigma_{\hat{p}} = 0.028$

Зад.7 Предполагаме, че популационната пропорция на децата от предишната задача p е всъщност 0.5. Каква е вероятността да наблюдаваме извадкова пропорция толкова голяма или по-голяма от отчетената $\hat{p} = 0.55$?

Отговор: $\mathbf{P}(\hat{p} > 0.55) = 0.0384$

Зад.8 Случайна извадка с големина n е избрана от биномна популация с популационен параметър p. Намерете средното и стандартното отклонение на извадковата пропорция $\hat{p}$, ако:

а) $n = 100, p = 0.3$; б) $n = 400, p = 0.1$; в) $n = 250, p = 0.6$.

Отговор: а) 0.3, 0.0458 б) 0.1, 0.015 в) 0.6, 0.0310

Зад.9 Начертайте всяко от извадковите разпределения от предишната задача. За всяко отбележете p и интервала $p \pm 2\sigma_{\hat{p}}$ на абсцисата $\hat{p}$. За подточка а) зачертайте площта под графиката, която е свързана с вероятността $\hat{p}$ да лежи в границите 0.08 около популационното средно и намерете тази вероятност.

Зад.10 Случайна извадка с големина $n = 500$ е избрана от биномна популация с популационен параметър $p = 0.1$.

а) Препоръчително ли е да се използва нормално разпределение за приближение на извадковото разпределение на $\hat{p}$? Проверете дали необходимите условия са изпълнени.

Използвайки резултатите от подточка а) и намерете вероятността, че:

б) $\hat{p} > 0.12$

в) $\hat{p} < 0.10$

г) $\hat{p}$ ще лежи в граници 0.02 около p.

Отговор: а) да б) 0.0681 в) 0.5 г) 0.8638

Зад.11 Средната заплата на преподаватели в щата Ню Йорк е \$29000, следвана от тази в окръг Колумбия - \$28621. Ако сме избрали случайна извадка от 40 преподаватели от щата Ню Йорк и 40 преподаватели от окръг Колумбия, каква е вероятността, че извадковата средна заплата $\bar{x}_1$ за Ню Йорк ще надвишава извадковата средна заплата $\bar{x}_2$ за окръг Колумбия с \$1000 или повече? (Предполагаме, че стандартното отклонение за разпределенията на двете популации заплати са приблизително $\sigma_1 = \sigma_2 = \$5000$)

Отговор: $\mathbf{P}(\bar{x}_1 - \bar{x}_2 \geq 1000) = 0.2877$

Зад.12 Независими сл. извадки от n_1 и n_2 наблюдения са избрани от популации със средно μ_1 и μ_2 и дисперсии σ_1^2 и σ_2^2. Намерете средното и стандартното отклонение на извадковото разпределение на разликата $(\bar{x}_1 - \bar{x}_2)$ на извадковите средни за:

а) $n_1 = 16, \mu_1 = 10, \sigma_1^2 = 4$ и $n_2 = 20, \mu_2 = 20, \sigma_2^2 = 8$;

б) $n_1 = 100, \mu_1 = 640, \sigma_1^2 = 1$ и $n_2 = 100, \mu_2 = 642, \sigma_2^2 = 3$;

Отговор: а) -10, 0.8062 б) -2, 2

Зад.13 Имаме условието от предишната задача.

а) Ако популациите са нормални, какво е разпределението на $(\bar{x}_1 - \bar{x}_2)$ за двете подточки?

б) Ако популациите не са нормални, какво можем да кажем за разпределението на $(\bar{x}_1 - \bar{x}_2)$ за двете подточки?

Отговор: а) нормално; б) приблизително нормално за б)

Зад.14 Независими сл. извадки от $n_1 = n_2 = 50$ наблюдения са избрани от популации с еднакви средни $\mu_1 = \mu_2$ и стандартни отклонения $\sigma_1 = 12$ и $\sigma_2 = 15$. Намерете:

а) $\mathbf{P}[(\bar{x}_1 - \bar{x}_2) \geq 3]$; б) $\mathbf{P}[|\bar{x}_1 - \bar{x}_2| \geq 3]$

Отговор: а) 0.1357 б) 0.2714

Зад.15 Независими случайни извадки с големини n_1 и n_2 са взети от биномни популации с параметри p_1 и p_2. Намерете средното и стандартното отклонение на извадковото разпределение на разликата $(\hat{p}_1 - \hat{p}_2)$ в извадковите пропорции за:

а) $n_1 = 100, p_1 = 0.5, n_2 = 300, p_2 = 0.4$ б) $n_1 = 400, p_1 = 0.1, n_2 = 400, p_2 = 0.6$

Отговор: а) 0.1, 0.0574 б) -0.5, 0.0287

Зад.16 С условието от предишната задача начертайте извадковото разпределение от подточка а). Отбележете средното и интервала $p_1 - p_2 \pm 2\sigma_{(\hat{p}_1 - \hat{p}_2)}$ по абсцисата $(\hat{p}_1 - \hat{p}_2)$. Защриховайте площта под кривата, която е свързана с вероятността разликата $(\hat{p}_1 - \hat{p}_2)$ да се отличава от $(p_1 - p_2)$ с по-малко от 0.06 и намерете тази вероятност.

Отговор: 0.7062

Зад.17 Независими случайни извадки с големини n_1 и n_2 са взети от биномни популации с параметри $p_1 = 0.3$ и $p_2 = 0.4$. Намерете стандартното отклонение на извадковото разпределение на разликата $(\hat{p}_1 - \hat{p}_2)$ в извадковите пропорции, ако:

а) $n_1 = 25, n_2 = 50$ б) $n_1 = 50, n_2 = 50$ в) $n_1 = 500, n_2 = 500$

Какъв е ефектът на увеличаването на големината на извадката върху стандартното отклонение на $(\hat{p}_1 - \hat{p}_2)$?

Отговор: а) 0.1149 б) 0.0949 в) 0.03

Зад.18 Един оператор в два супермаркета открил, че загубите от продажбите поради кражби, изтекъл срок на годност и др. в супермаркет 1 се менят със средно 1237 долара на ден и стандартно отклонение 183 долара на ден. Същите дневни загуби в супермаркет 2 са със средно 1485 долара и стандартно отклонение 259 долара на ден. На колко е равна вероятността за това, че общата дневна загуба в двата супермаркета може да надхвърли 3000 долара. Предполагаме, че дневните загуби в двата супермаркета са приблизително нормални.

Отговор: 0.1894