ВОЙНА УЙОБОК

К/р МатСтат

Непараметричне оцінювання.

Приклад 1

Завдання 1. Неперервна випадкова величина ξ представляє результат вибору точки на відрізку [2, 3]. Щільність f(x) випадкової величини ξ визначається рівністю:

$$f_\xi(x) = \begin{cases} 0, \text{ якщо } x < 2, \\ a \cdot x, \text{ якщо } 2 \le x \le 3 \\ 0, \text{ якщо } x > 3 \end{cases}.$$

Обчислити ймовірність того, що точка , вибрана на відрізку [2, 3], потрапить до інтервалу [2,25 < ξ < 2,75]. Виконати всі необхідні для цього обчислення. Привести графічну ілюстрацію.

Скорист. власт. ф-ї щільності:

$\int_{-\infty}^{\infty} f(x)dx = 1$. Тоді маємо: $\int_{-\infty}^{2} 0dx + \int_{2}^{3} axdx + \int_{3}^{\infty} 0dx =$

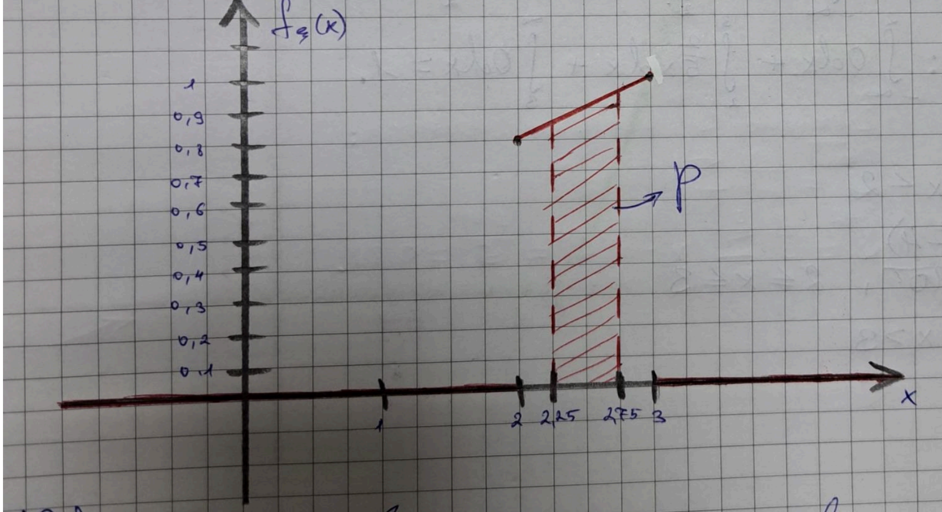
$$= a\int_{2}^{3} xdx = a\cdot \frac{x^2}{2}\Big|_{2}^{3} = a\left(\frac{9}{2}-\frac{4}{2}\right) = a\cdot\frac{5}{2} = 1 \Rightarrow a = \frac{2}{5}.$$

Скорист. ф-лою:

$$P\{c<\xi<d\} = \int_{c}^{d} f(x)dx$$

$$P\{2,5<\xi<2,75\} = \int_{2,25}^{2,75}\frac{2}{5}xdx = \frac{2}{5}\cdot\frac{x^2}{2}\Big|_{2,25}^{2,75} = \frac{2}{5}\cdot\left(\frac{7,5625-5,0625}{2}\right) =$$

$$= \frac{2}{5}\cdot 1,25 = 0,5.$$



$P\{2,25<\xi<2,75\} = p$ – площа заштрихованої трапеції.

Завдання 2. Навести визначення дисперсії випадкової величини. Обчислити дисперсію $\sigma_\xi^2 = D(\xi)$ виконавши всі необхідні для цього обчислення.

Визначення. Дисперсією випадкової величини ξ називається математичне сподівання квадрату відхилення випадкової величини ξ від її математичного сподівання, тобто число $D(\xi)$, що визначається наступним чином:

$$D(\xi) = E(\xi - E(\xi))^2.$$

Обчисл. $E(\xi^2)$ за ф-лою:

$$E(\xi^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx.$$

$$E(\xi^2) = \frac{2}{5}\int_2^3 x^3\,dx = \frac{2}{5}\cdot\frac{x^4}{4}\Big|_2^3 = \frac{2}{5}\left(\frac{81-16}{4}\right) = 6,5$$

Обчисл. $m = E(\xi)$ за ф-лою:

$$E(\xi) = \int_{-\infty}^{\infty} x f(x)\,dx.$$

$$E(\xi) = \frac{2}{5}\int_2^3 x^2\,dx = \frac{2}{5}\cdot\frac{x^3}{3}\Big|_2^3 = \frac{2}{5}\left(\frac{27}{3} - \frac{8}{3}\right) = 2,53$$

$$\delta^2 = E(\xi^2) - (E(\xi))^2 = 6,5 - 2,53^2 = 6,5 - 6,4009 = 0,0991.$$

Завдання 3. Привести визначення неперервної випадкової величини. Використовуючи його знайти функцію розподілу $F_\xi(x)$ випадкової величини ξ, що має щільність $f_\xi(x)$.

Визначення. Кажемо, що ξ є *неперервною* випадковою величиною з *щільністю* $f_\xi(x)$, якщо її функцію розподілу $F_\xi(x) = P\{\xi < x\}$ для будь-якого дійсного числа x можна записати у вигляді:

$$F_\xi(x) = \int_{-\infty}^{x} f_\xi(u)\,du, \quad -\infty < x < \infty.$$

$F(x) = \int_{-\infty}^{x} f_\xi(x)\,dx$ – з визначення.

При $x < 2$: $\int_{-\infty}^{x} 0\,dx = 0$

При $2 \le x \le 3$: $\int_{-\infty}^{2} 0\,dx + \int_2^x \frac{2}{5}x\,dx = \frac{2}{5}\int_2^x x\,dx = \frac{2}{5}\cdot\frac{x^2}{2}\Big|_2^x = \frac{2(x^2-4)}{10}.$

При $x > 3$: $\int_{-\infty}^{2} 0\,dx + \int_2^3 \frac{2}{5}x\,dx + \int_3^{\infty} 0\,dx = 1.$

$$F(x) = \begin{cases} 0, & x < 2 \\ \dfrac{2(x^2-4)}{10}, & 2 \le x \le 3 \\ 1, & x > 3 \end{cases}$$

Завдання 4. З популяції, функція розподілу $F_\xi(x)$ якої була визначена в завданні 3, вибрана випадкова вибірка:

x1 = 2,1; x2 = 2,7; x3 = 2,9; x4 = 2,2; x5 = 2,3; x6 = 2,7; x7 = 2,1; x8 = 2,7; x9 = 2,1; x10 = 2,5.

Пояснити, що називається«варіаційним рядом x*1, x*2, ..., побудованим на підставі вибірки {x1, x2, ..., xn}». Побудувати варіаційний ряд, використовуючи дану випадкову вибірку, та знайти значення функції розподілу $F_\xi(x)$ в точках x^*_4 та x^*_9 цього ряду.

Нехай $\{\xi_1, \xi_2, \dots, \xi_n\}$буде простою вибіркою, вибраною з популяції, що має функцію розподілу $F(x)$. Розташуємо її елементи в зростаючому порядку:

$$\xi^*_1 \le \xi^*_2 \le \dots \le \xi^*_n.$$

В математичній статистиці отриманий таким чином вектор $\{\xi^*_1, \xi^*_2, \dots, \xi^*_n\}$ називається *варіаційним рядом*.

Побуд. варіаційний ряд:

$x_1^* = 2,1$, $x_2^* = 2,1$, $x_3^* = 2,1$, $x_4^* = 2,2$, $x_5^* = 2,3$, $x_6^* = 2,5$, $x_7^* = 2,7$, $x_8^* = 2,7$, $x_9^* = 2,7$, $x_{10}^* = 2,9$.

$$F_\xi(x_4^*) = F(2,2) = \frac{2(2,2^2-4)}{10} = 0,168$$

$$F_\xi(x_9^*) = F(2,7) = \frac{2(2,7^2-4)}{10} = 0,658$$

Завдання 5. Привести визначення емпіричної функції розподілу, побудованої на підставі статистичної вибірки {x1, x2, ..., xn}.

Нехай $\hat{F}_n(x)$- емпірична функція розподілу, побудована на підставі даної вибірки. Знайти значення емпіричної функції розподілу $\hat{F}_n(x)$ в точках x^*_4 та x^*_9 варіаційного ряду.

Визначення. *Емпіричною функцією розподілу*, побудованою на підставі простої статистичної вибірки $(\xi_1, \xi_2, \dots, \xi_n)$ називається функція

$$\hat{F}_n(x), -\infty < x < +\infty,$$

яка визначає розподіл значення ознаки X в цій вибірці.

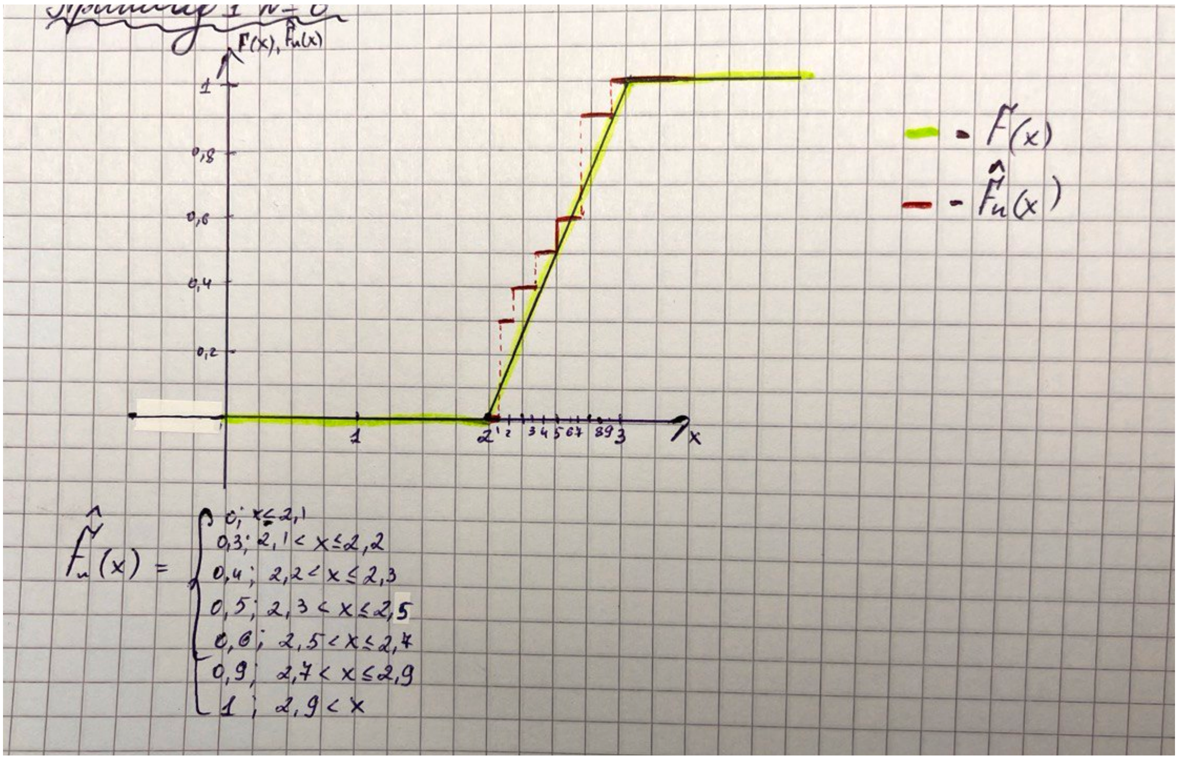
Іншими словами, для кожного дійсного числа x $\hat{F}_n(x)$ вказує частоту появи у вибірці $(\xi_1, \xi_2, \dots, \xi_n)$ чисел менших від x.

$\hat{F}_n(x) = \frac{D_n(x)}{n}$, де $D_n(x)$ – к-сть елементів, що передують x у варіаційному ряді.

$$\hat{F}_n(x_4^*) = \hat{F}_n(2,2) = \frac{3}{10} = 0,3$$

$$\hat{F}_n(x_9^*) = \frac{8}{10} = 0,8$$

Завдання 6. Побудувати на тому самому малюнку графіки функції розподілу F(x) та емпіричної функції розподілу $\hat{F}_n(x)$



(Червона лінія емпір. функції проходить ще по осі х до 2,1)

також та жовта хуйня має йти не тільки до нуля, а по всій осі де х<0 lol

Приклад 2. (5 балів)

Припустимо, що функція розподілу F(x) генеральної популяції визначається формулою:

$$F(x) = \begin{cases} 0, dla \ x < -1 \\ \dfrac{x+1}{5}, \ dla \ -1 \le x \le 4. \\ 1, \ dla \ x > 4 \end{cases}$$

Результатом спостереження неперервної випадкової величини, що має функцію розподілу F(x), є наступна проста вибірка:

3,781; 2,296; –0,41; 2,100; 0,174; 0,300; 0,875; 1,815; 1,922; 2,546; 1,166; 1,348; 2,623;

0,987; 4; 0,205; 2,103; 3,619; 0,065; 2,377; –0,881; –0,498; –1; 2,723; 2,816; 2,485; 3,054;

1,851; 3,744; 1,336; 1,882; 3,379; 2,673; 1,300; –0,742; 0,001; 3,026; 0,298; 0,609; 2,255;

0,195; 3,401; 3,219; 1,41; 3,522; 0,715; 0,301; 0,370; –0,776; 1,375.

($n = 50$, $x_{min} = -1$; $x_{max} = 4$)

Завдання 1.. Нехай $\hat{F}_n(x)$ – емпірична функція розподілу, побудована на підставі даної вибірки. Використовуючи визначення емпіричної функції розподілу $\hat{F}_n(x)$ знайти її значення в точках:

1) x1 = –0,95; 2) x2 = 1,5; 3) x3 = 2.

проводячи та пояснюючи всі необхідні для цього обчислення:

1) $x_1 = -0{,}95$; $\hat{F}_n(x_1)=\hat{F}_n(-0{,}95)=$....

2) $x_2 = 1{,}5$; $\hat{F}_n(x_2)=\hat{F}_n(1{,}5)=$....

3) $x_3 = 2$; $\hat{F}_n(x_3)=\hat{F}_n(2)=$....

1) $x_1 = -0{,}95$, $\hat{F}_n(x_1) = \hat{F}_n(-0{,}95) = \dfrac{\nu_n(-0{,}95)}{50} = \dfrac{1}{50} = 0{,}02$

2) $x_2 = 1{,}5$, $\hat{F}_n(x_2) = \hat{F}_n(1{,}5) = \dfrac{\nu_n(1{,}5)}{50} = \dfrac{25}{50} = 0{,}5$

3) $x_3 = 2$, $\hat{F}_n(x_3) = \hat{F}_n(2) = \dfrac{\nu_n(2)}{50} = \dfrac{29}{50} = 0{,}58$

Завдання 2. Використовуючи визначення неперервної випадкової величини знайти щільність f(x) генеральної популяції, що має задану в Прикладі 2 функцію розподілу F(x). Дати обґрунтовану відповідь на наступні питання:

$$f_\xi(x) = (F_\xi(x))'$$

$$f_\xi(x) = \begin{cases} 0, & x < -1 \\ \dfrac{1}{5}, & -1 \le x \le 4 \\ 0, & x > 4 \end{cases}$$

Завдання 3. Яким чином можна оцінити щільність f(x) в інтервалах $x\in[1; 2]$, та $x\in[3; 4]$,використовуючи приведену випадкову вибірку? Тобто пояснити, як виглядає функція, яку можна вважати наближенням щільності f(x) в цих інтервалах: $x\in[1; 2]$, та $x\in[3; 4]$

Щільність f(x) в інтервалах [1; 2] та [3; 4] можна оцінити як висоту прямокутників гістограми, де висота обчисл. за ф-лою:

$$h_i = \frac{n_i}{n \cdot (a_i - a_{i-1})}, \; i = 1, 2, ..., k.$$

Це дає наближення щільності на кожному інтервалі, що відповідає частоті значень у вибірці.

Завдання 4. Привести необхідні визначення та обчислювальні формули

- У випадку неперервного розподілу генеральної популяції проблему непараметричного оцінювання вирішує. гістограма $h(x)$, $-\infty < x < \infty$, побудована на основі простої вибірки $\{\xi_1, \xi_2, \dots, \xi_n\}$,

Визначення. Гістограмою побудованою на основі вибірки $\{\xi_1, \xi_2, \dots, \xi_n\}$ називається функція $h(x)$, $-\infty < x < \infty$, що визначається наступним чином:

$$h(x) = \begin{cases} 0, \; якщо \;\; x < a_0, \\ \dfrac{n_i}{n \cdot (a_i - a_{i-1})}, \; якщо \;\; x \in [a_{i-1}, a_i), \quad i = 1, 2, ..., k. \\ 0, \; якщо \; x \geq a_k \end{cases}$$

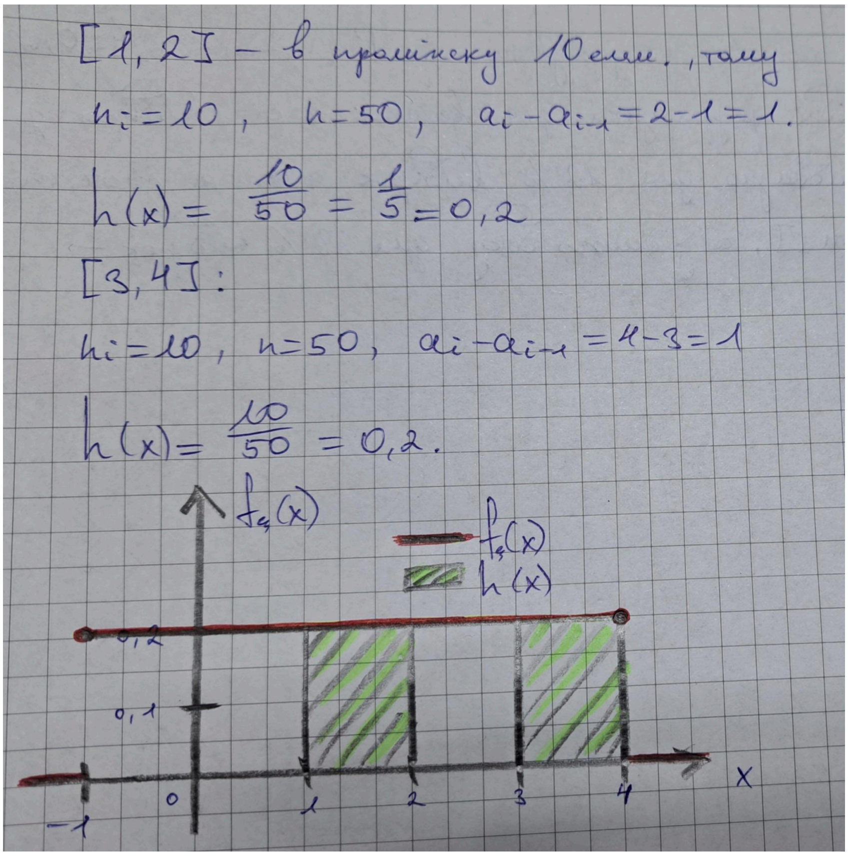
Графік функції $h(x)$ представляє верхню межу фігури, яка складається з прямокутників. Основою кожного окремого прямокутника є інтервал $[a_{i-1}, a_i]$, а його висота h_i – пропорційна чисельності n_i цього інтервалу і визначається формулою:

$$h_i = \frac{n_i}{n \cdot (a_i - a_{i-1})}, \; i = 1, 2, ..., k.$$

допис. до h(x)

a0 = -1, ak = 4, [ai-1; ai] = [1, 2] або [3, 4].

Завдання 5. Знайти значення функції, яку можна вважати наближенням щільності f(x). Побудувати на тому самому малюнку графіки функцій f(x) та її наближення в інтервалах $x \in [1; 2]$, та $x \in [3; 4]$



Надійні інтервали для параметрів нормального розподілу

Приклад 3. (5 балів)

Маємо наступні результати вимірювання невідомої величини m:

$x1 = 1{,}2$; $x2 = 0{,}6$; $x3 = 1{,}3$; $x4 = 2{,}3$; $x5 = 0{,}9$; $x6 = 3{,}1$; $x7 = 2{,}1$; $x8 = 1{,}8$; $x9 = 2{,}9$.

Вимірювання неточні, а похибку можна вважати випадковою величиною, яка має нормальний розподіл з параметрами $(0,\sigma^2)$. При цьому $\boldsymbol{\sigma^2=0.49}$

Завдання 1. Привести формулу емпіричного естиматора $\overline{m}$ для оцінювання величини **m**. Знайти оцінку (m*) невідомої величини m, використовуючи цю вибірку та емпіричний естиматор $\overline{m}$.

Отже статистикою для оцінювання параметру p (*естиматором*) є певна функція вибірки $p^*(\xi_1, \xi_2, \dots, \xi_n)$, за допомогою якої можемо судити про ймовірне значення цього параметру спираючись на проведені спостереження. Оскільки в математичній статистиці використовуємо випадкову вибірку, то статистика $p^*(\xi_1, \xi_2, \dots, \xi_n)$ для оцінювання параметру p також буде випадковою величиною.

Ця оцінка позначається символом $\overline{m}$ і називається *емпіричним середнім*, або значенням параметру $m = E(\xi)$ у вибірці. $\overline{m}$ є статистичною оцінкою математичного сподівання: $\overline{m} \approx m$.

$$\overline{m} = \frac{1}{n}\sum_{i=1}^{n}\xi_i \ .$$

Знайдемо m* використовуючи емпіричний естиматор:

$$\overline{m} = \sum_{i=1}^{n} \xi_i$$

m*= $\frac{1}{9}\sum_{i=1}^{n} x_i = \frac{1}{9}(1.2+0.6+1.3+2.3+0.9+3.1+2.1+1.8+2.9)= \frac{1}{9}(16.2)=1.8$

Завдання 2. Що означає вираз: «[u1, u2] – надійний інтервал для параметру m з рівнем надійності γ = 0,88»?

Визначення: Надійним інтервалом для параметру p, з рівнем надійності γ називаємо будь-який випадковий інтервал $[u_1, u_2]$, що задовольняє наступні умови:

$$\mathrm{P}\{u_1 < u_2\} = 1; \ \mathrm{P}\{u_1 < p < u_2\} = \gamma.$$

Довжина $(u_2 - u_1)$ цього інтервалу також буде випадковою величиною і може бути ужита до визначення *точності оцінювання* параметру p. Для цього достатньо припустити, що *істинне* значення параметру p знаходиться в інтервалі $[u_1, u_2]$.

Згідно з визначенням надійного інтервалу таке припущення дуже ймовірне, оскільки $\gamma \approx 1$. Існує однак ризик, що інтервал $[u_1, u_2]$ не містить в собі значення параметру p. Його величина (α) дорівнює:

$$\alpha = 1 - \gamma.$$

Встановлюючи рівень надійності γ особа, що приймає рішення, тим самим визначає допустимий для неї рівень ризику (α).

Це означає, що ми можемо стверджувати, що $P\{u_1 < m < u_2\} = 0,88$. Якщо проводити подібні розрахунки для великої к-сті вибірок з цієї популяції, то приблизно для 12 % вибірок фактичне знач. m не потрапить до $[u_1, u_2]$, а з іншого боку для 88 % вибірок – потрапить.

Завдання 3. Привести формули для обчислення меж u_1 та u_2 цього інтервал та пояснити, які з величин в цих формулах безпосередньо залежать від γ

Оскільки за умовою дисперсія є відомою використовуємо нормальний розподіл.

- Нехай γ – визначений в умові задачі рівень надійності.

$$\alpha = 1 - \gamma.$$

Для цього використовуючи функцію Лапласа $\Phi(x)$ знаходимо z_α з рівності:

$$\Phi(z_\alpha) = P\{N(0,\ 1) < z_\alpha\} = 1 - \frac{\alpha}{2}.$$

$$P\left\{\overline{m} - z_\alpha \frac{\sigma}{\sqrt{n}} < m < \overline{m} - z_\alpha \frac{\sigma}{\sqrt{n}}\right\} = \gamma. \ u_1 = \overline{m} - z_\alpha \frac{\sigma}{\sqrt{n}}; \ u_2 = \overline{m} + z_\alpha \frac{\sigma}{\sqrt{n}}.$$

При знаходженні u1, u2 використовуємо z_a яке залежатиме від γ

Завдання 4. Пояснити, яким чином визначити значення вказаних в попередньому завданні величин. **Привести графічну ілюстрацію**

$$\overline{m} = \frac{1}{n}\sum_{i=1}^{n} \xi_i .$$

- Нехай γ – визначений в умові задачі рівень надійності.

$$\alpha = 1 - \gamma.$$

Для цього використовуючи функцію Лапласа $\Phi(x)$ знаходимо z_α з рівності:

$$\Phi(z_\alpha) = P\{N(0,\ 1) < z_\alpha\} = 1 - \frac{\alpha}{2} .$$

σ=0,7 з умови
n= 9 - обсяг вибірки

Завдання 5. Побудувати надійний інтервал [u1, u2], використовуючи приведену вибірку та знайдені оцінки параметрів.

$\Phi(z_a) = P(N(0, 1) < z_a) =$ $\boxed{1-\frac{1-y}{2}=0.94}$

$\Phi(1.55) = 0.9394 < \Phi(z_a) = 0.94 < \Phi(1.56) = 0.9406$

$z_a = 1.555$, o = 0.7, $\boxed{\overline{m}=1.8}$

$$\boxed{\delta=\frac{z_a * o}{\sqrt{n}}=\frac{1.555*0.7}{3}=0..363}$$

$u_1 = 1.8 – 0.363 = 1.437$; $u_1 = 1.8 + 0.363 = 2.163$;

Завдання 6. Пояснити, яким чином використовуючи надійний інтервал $[u_1,u_2]$ можна оцінити точність (**δ**) наближення за допомогою оцінки (**m***) невідомого значення **m** величини, яка вимірюється А саме:

- Записати відповідну формулу для обчислення мінімального значення **δ** , що відповідає умові:

$$|m - m^*| \leq \delta$$

- Знайти величину **δ** для приведеної вибірки.

$$|m - m^*| < \boxed{\delta = \frac{z_a * o}{\sqrt{n}}}\ , \quad \delta = 0.363$$

В душе не єбу, що означає пояснити. Дав відповідь на питання. Щоб не писати трильйон формул просто порахував на калькуляторі відповідь

Завдання 7. Дати обґрунтовану відповідь на наступне питання:

- Якою буде надійність знайденої точності **δ** ?Тобто знайти для отриманого в попередньому завданні значення δ величину (p) ймовірності наступної випадкової події:

$$p = P\{|m - m^*| \leq \delta\}$$

|m – m*| являє собою відстань від значення m до центру m* інтервалу $[u_1, u_2]$, Тобто |m – m*| <= δ це потрапляння значення m до інтервалу $[u_1, u_2]$, а тоді ймовірність
P(|m – m*| <= δ) = p = y, оскільки y це і є ймовірність потрапляння значення m до інтервалу $[u_1, u_2]$. Тобто p = y = 0.88

Завдання 8. Припустимо тепер, що інтервал необхідно побудувати таким чином, щоб значення ймовірності **p** дорівнювало: **p = 0,98**.

- Якою буде в цьому випадку точність (**δ**) наближення за допомогою оцінки (**m***) невідомої величини математичного сподівання (m) генеральної популяції.

Тобто пояснити, **які зміни** в попередніх обчисленнях необхідно провести та **виконати** ці зміни.

p = y = 0.98, $\boxed{\delta = \frac{z_a * o}{\sqrt{n}}}$, o = 0.7, n = 9

Необхідно знайти z_a:

a = 1 - y = 0.02 => $P(|N(0,1)| < z_a) = 1 - a = 0.02$

$\Phi(z_a) = P(N(0,1) < z_a) = \boxed{1 - \frac{a}{2} = 0.99}$

$\Phi(2.32) = 0.9898 < \Phi(z_a) = 0.99 < \Phi(2.33) = 0.9901 \Rightarrow z_a = 2.325$

$$\boxed{\delta = \frac{2.325 * 0.7}{3} = 0.5425}$$

Критерії істотності

Приклад 4. (5 балів) Припустимо, що математичне сподівання **m** популяції, яка має нормальний розподіл, відоме **m=6**. Для оцінки невідомої дисперсії $\sigma^2=D(\xi)$ маємо просту вибірку: $(\xi_1;\xi_2;\xi_3;\xi_4;\xi_5)$. На підставі цієї вибірки необхідно перевірити наступну параметричну гіпотезу: **H: σ^2=2,5**.

Завдання 1. Припустимо, що ймовірність відхилення вірної гіпотези не повинна перевищити **0,03** обчислити. Обчислити верхню **(g)** та нижню **(d)** межі інтервалу [**d**;**g**] відповідного критерію істотності. Обґрунтувати відповідь тобто:

а) Привести відповідні формули

б) Пояснити, яким чином визначаємо необхідні для обчислень величини. Привести та пояснити їх графічну ілюстрацію.

а) Межі інтервалу визначаються такими формулами:

$$d = \frac{h_{1\alpha}}{n} \cdot \sigma_0^2$$
$$g = \frac{h_{2\alpha}}{n} \cdot \sigma_0^2$$
(лек 16а)

б) Числа $h_{1\alpha}$ та $h_{2\alpha}$ визнач. з рівностей:

$$P\{\chi^2_{(n)} < h_{1\alpha}\} = \frac{\alpha}{2}, \text{ натомість: } P\{\chi^2_{(n)} < h_{2\alpha}\} = 1 - \frac{\alpha}{2}.$$ (лек 15)

, де a = 0,03 - рівень істотності;

σ_0^2- гіпотетична дисперсія, $\sigma_0^2 = 2,5$;

n -розмір вибірки, n = 5.

Знайдемо $h_{1\alpha}$ та $h_{2\alpha}$ використовуючи таблицю $\chi^2_{(5)}$- розподілу:

$P\{\chi^2_{(5)} < h_{1\alpha}\} = 0,03/2 = 0,015.$

З таблиці:

$H_5(0,66) < H_5(h_{1\alpha}) < H_5(0,67)$, де $H_5(x)$ - функція розподілу $\chi^2_{(5)}$.

Врах. монотонність:

$0,66 < h_{1\alpha} < 0,67$

$h_{1\alpha} \sim 0,66$

$P\{\chi^2_{(5)} < h_{2\alpha}\} = 1 - 0,03/2 = 0,985.$

З таблиці:

$H_5(14,1) = 0,985 \Rightarrow h_{2\alpha} = 14,1.$

$$d = \frac{h_{1\alpha}}{n} \cdot \sigma_0^2 = 0,66/5 * 2,5 = 0,33;$$

$$g = \frac{h_{2\alpha}}{n} \cdot \sigma_0^2 = 14,1/5 * 2,5 = 7,05.$$

Завдання 2. Охарактеризувати статистику критерію істотності для перевірки гіпотези **H: σ^2=2,5**, що буде застосовуватись в ***даному прикладі***, тобто:

(а). Привести формулу статистики.

(б). Вказати розподіл статистики та параметри цього розподілу у випадку, якщо гіпотеза вірна.

а) В якості статистики обираємо емпіричну дисперсію $S^2_{(m)}$:

$S_{(m)}^2 = \frac{1}{n}\sum_{i=1}^{n}(\xi_i - m)^2$.

б) $S_{(m)}^2 \Leftrightarrow \frac{\sigma^2}{n} * \chi_{(n)}^2$, де $\chi_{(n)}^2$ випадкова величина, яка має хі-квадрат розподіл з n степенями свободи.

Завдання 3. Припустимо, що в результаті проведених спостережень отримано конкретні значення (x1,x2,x3,x4,x5) для вибірки **($\xi_1;\xi_2;\xi_3;\xi_4;\xi_5$)**, а обчислене на їх підставі значення статистики ($S^2_{(m)}$) критерію істотності дорівнює: **$S^{*2}_{(m)}$=0,55**

Яке рішення щодо статистичної гіпотези **H: σ^2=2,5**. буде прийнято? Чому? Обґрунтувати його, будуючи **критичну область Q_α**

Що означає параметр **α**, та яке його числове значення?

Критична область Q_α має вигляд:

Q_α = [0; d] ∪ [g; +∞)

Q_α = [0; 0,33] ∪ [7,05; +∞)

Оскільки d < $S_{(m)}^2$ < g, тобто знач. $S_{(m)}^2$ не потрапило до критичної обл. Q_α, то гіпотезу приймаємо.

α - рівень істотності, тобто ймовірність припустити помилку 1-го роду - відхилити вірну гіпотезу: α = 0,03.

Завдання 4. Яким чином зміниться побудований в завданні 1 критерій якщо зміниться гіпотеза **H: σ^2=2**, а рівень істотності **α=0,03** не зміниться? Яке рішення буде прийнято в цьому випадку? **Обґрунтувати його**.

Оскільки рівень істотності α = 0,03 не змінився, то $h_{1\alpha}$ та $h_{2\alpha}$ залишаються тими самими. Перерахуємо d та g:

(σ_0^2 = 2, n = 5)

$$d = \frac{h_{1\alpha}}{n}\cdot\sigma_0^2 \quad = 0,33*2/5 = 0,132$$

$$g = \frac{h_{2\alpha}}{n}\cdot\sigma_0^2 \quad = 14,1*2/5 = 5,64$$

Критична обл. Q_α = [0; d] ∪ [g; +∞)

Q_α = [0; 0,132] ∪ [5,64; +∞).

Оскільки d < $S_{(m)}^{2}$ < g, тобто знач. $S_{(m)}^{2}$ не потрапило до критичної обл. Q_{α}, то гіпотезу приймаємо.

Завдання 5. Яке рішення буде прийнято, якщо ***зміняться*** як гіпотеза: **H:σ²=2**, так і рівень істотності **α=0,23**?

Оскільки рівень істотності α = 0,23 змінився, то шукаємо знач. $h_{1\alpha}$ та $h_{2\alpha}$:

$H_5(h_{1\alpha})$ = α/2 = 0,23/2 = 1,115.

З таблиці:

$H_5(1{,}7) < H_5(h_{1\alpha}) < H_5(1{,}71)$.

Врах. монотонність:

$1{,}7 < h_{1\alpha} < 1{,}71$

$h_{1\alpha} \sim 1{,}7$.

$H_5(h_{2\alpha})$ = 1 - α/2 = 1 - 0,23/2 = 0,885.

З таблиці:

$H_5(8{,}85) < H_5(h_{2\alpha}) < H_5(8{,}86)$.

Врах. монотонність H_5(x):

$8{,}85 < h_{1\alpha} < 8{,}86$

$h_{1\alpha} \sim 8{,}85$.

Обчислимо d та g:

$$d = \frac{h_{1\alpha}}{n} \cdot \sigma_0^2 \quad = 1{,}7*2/5 = 0{,}68;$$

$$g = \frac{h_{2\alpha}}{n} \cdot \sigma_0^2 \quad = 8{,}85*2/5 = 3{,}54.$$

Критична обл. Q_{α} = [0; d] ◡ [g; +∞)

Q_{α} = [0; 0,68] ◡ [3,54; +∞).

Оскільки знач. $S_{(m)}^{2}$ є [0; 0,68], тобто знач. $S_{(m)}^{2}$ потрапило до критичної обл. Q_{α}, то гіпотезу відхиляємо.

Завдання 2;5 з ср5 Бреславець

Дописати приклад 4 1б