Практичне завдання 1.

Задачі перевірки статистичних гіпотез

1.1. Критерії, функції потужності критеріїв

Задача перевірки статистичних гіпотез. Часто виникає необхідність у розв'язанні такої задачі: маємо стохастичний експеримент, що полягає в спостереженні випадкової величини $\xi = (\xi_1,\xi_2,\ldots,\xi_n)$ зі значеннями в просторі R^n або його частині, тобто в одержанні вибірки обсягом n. (Ми будемо розглядати такі стохастичні експерименти, які можна описати моделями типу $\{R^n, \Re^n, P_\theta\}$, $P_\theta \in \wp$, $\wp = \{P_\theta\cdot, \theta \in R^s\}$, де $\wp$ – деяка, як у нас, параметрична сукупність розподілів). Щодо розподілу випадкової величини $\xi = (\xi_1,\xi_2,\ldots,\xi_n)$ (стосовно розподілу вибірки ξ) відомо тільки те, що він належить до класу $\wp$. З цього класу вибираємо (з тих чи інших міркувань) один із розподілів, наприклад G, і як модель випадкової величини $\xi = (\xi_1,\xi_2,\ldots,\xi_n)$ пропонуємо розглядати $\{R^n, \Re^n, G\}$.

Мета полягає в тому, щоб за результатом $\xi = (\xi_1,\xi_2,\ldots,\xi_n)$ експерименту (відомий він чи одержимо його потім) дійти висновку: експеримент може описуватися моделлю $\{R^n, \Re^n, G\}$ {G може бути розподілом випадкової величини ξ) або експеримент не може описуватися моделлю $\{R^n, \Re^n, G\}$ (G не може бути розподілом випадкової величини ξ). Зробити за результатом експерименту "сильніший" висновок – (G є розподілом випадкової величини ξ – не можна, оскільки один і той самий результат можливий при різних розподілах, а не тільки при істинному.

У теорії перевірки статистичних гіпотез прийнято такі означення та домовленості.

Гіпотези щодо розподілів випадкових величин називають *статистичними*.

Вибір розподілу (чи класу розподілів) із сукупності $\wp$ будемо називати вибором *основної (нульової) гіпотези* щодо розподілу в моделі $\{R^n, \Re^n, P_\theta\}$ (щодо розподілу випадкової величини ξ зі значеннями в просторі R^n). Після вибору основної гіпотези решту гіпотез називають *альтернативними або конкуруючими* відносно основної (нульової). Нульову гіпотезу будемо позначати H_0. Сукупність конкуруючих (альтернативних) гіпотез, як правило, буде параметричною. Конкуруючі гіпотези будемо позначати так: H_θ, $\theta \in \Theta$.

Розглянемо, наприклад, біноміально розподілену випадкову величину, що набуває значень $0,1,\ldots,n$; n - відоме. Розподілом ξ може бути будь-який розподіл із класу $\wp = \{P_\theta\cdot, \theta \in (0,1)\}$, де

$$P_\theta(k) = C_n^k \theta^k (1-\theta)^{n-k},\ k = 0,1,\ldots,n;\ \theta \in \Theta. \quad (1)$$

Вибір із сукупності $\wp$ одного з розподілів, наприклад

$$P_{1/2}(k) = C_n^k\left(\frac{1}{2}\right)^k\left(1-\frac{1}{2}\right)^{n-k},\ k= 0,1,...,n, \qquad (2)$$

це вибір нульової (основної) гіпотези щодо розподілу випадкової величини ξ. Після того як основну гіпотезу вибрано (у розглядуваному випадку це гіпотеза: ξ має розподіл $P_{1/2}(k)$, $k= 0,1,...,n$). інші гіпотези стають конкуруючими відносно H_0 (сукупністю конкуруючих гіпотез є H_θ, $\theta\in(0,1/2)\cup(1/2,1)$, тобто розподілом ξ є $P_\theta(k)$ $k= 0,1,...,n$; $\theta\neq1/2$.

Гіпотези щодо розподілів, які однозначно їх визначають, будемо називати *простими*, у противному разі - *складними*.

Наприклад, гіпотеза H_0: випадкова величина ξ має розподіл

$$P_{\theta_0}(k) = C_n^k\theta_0^k(1-\theta_0)^{n-k},\ k= 0,1,...,n, \qquad (3)$$

де θ_0 - фіксоване, є простою. Гіпотеза H_θ: випадкова величина ξ має розподіл

$$P_\theta(k) = C_n^k\theta^k(1-\theta)^{n-k},\ k= 0,1,...,n, \qquad (4)$$

де $\theta\in(1/4;\ 1/2)$, є складною.

Формулюючи задачу перевірки статистичних гіпотез, за нульову гіпотезу (для наочності та простоти) ми взяли просту гіпотезу: розподілом випадкової величини ξ є G, де G – цілком визначений розподіл.

Зазначимо, що математична статистика не дає ніяких рекомендацій щодо вибору нульової гіпотези, який повністю визначається дослідником і залежить від поставленої задачі.

Методика перевірки статистичних гіпотез. Нехай основна гіпотеза H_0 полягає в тому, що розподілом випадкової величини ξ зі значеннями в просторі R^n є розподіл G. Необхідно перевірити гіпотезу H_0, тобто дійти висновку: G може бути розподілом випадкової величини ξ (будемо говорити "гіпотеза H_0 не відхиляється") або G не може бути розподілом випадкової величини ξ (будемо говорити "гіпотеза H_0 відхиляється"). Робити висновок про відхилення або невідхилення гіпотези H_0: розподілом випадкової величини ξ є G ми будемо за результатом $\xi(\omega)$ експерименту, за реалізацією $\xi(\omega)$ вибірки ξ (крім реалізації $\xi(\omega)$ вибірки ξ ми не маємо нічого, що несло б інформацію про її розподіл). Тому, щоб за результатом експерименту (за реалізацією $\xi(\omega)$ вибірки ξ) зробити висновок про відхилення або невідхилення гіпотези H_0, необхідно вказати ті результати $\xi(\omega)=(\xi_1(\omega),\xi_2(\omega),\ldots,\xi n(\omega))$ експерименту, при яких ми будемо відхиляти H_0, і ті, при яких ми не будемо відхиляти H_0. Інакше кажучи, треба поділити множину можливих результатів $\xi(\omega)=(\xi_1(\omega),\xi_2(\omega),\ldots,\xi n(\omega))$ (вибірковий простір) на дві неперетинні множини: S – множину результатів, при яких H_0 відхиляється, та $\overline{S}$, – при яких H_0 не відхиляється. Далі проводимо експеримент – одержуємо реалізацію $\xi(\omega)$ = $\xi(\omega)=(\xi_1(\omega),\xi_2(\omega),\ldots,\xi n(\omega))$ випадкової величини $\xi = (\xi_1,\xi_2,\ldots,\xi_n)$, тобто вибірку. Якщо при цьому $\xi(\omega)$ потрапляє до S, то гіпотезу H_0 відхиляємо, у противному разі – не відхиляємо.

Ось такою і є методика перевірки статистичних гіпотез.

Означення. Борелеву множину S таку, що при $\xi\in S$ гіпотеза H_0 відхиляється, а при $\xi\notin S$ – не відхиляється, будемо називати критичною множиною (критичною областю) або критерієм для перевірки гіпотези H_0.

Визначення в R^n – вибірковому просторі – борелевої множини S рівнозначне визначенню борелевої функції $\varphi(x) = I_S(x)$ – індикатора множини S:

$$I_S(x) = \varphi(x) = \begin{cases} 1, \text{ якщо } x\in S, \\ 0, \text{ якщо } x\notin S. \end{cases} \qquad (5)$$

І оскільки між множинами S та їхніми індикаторами $I_S(x)$ існує взаємно однозначна відповідність, борелеву функцію $\varphi(x)=I_S(x)$ також називають *критерієм* (іноді тестом) для перевірки гіпотези H_0. Якщо $\varphi(x)=1$ (або, що те саме, $\xi\in S$), то гіпотеза H_0 відхиляється, а якщо $\varphi(x)=0$ (або, що те саме, $\xi\notin S$), то гіпотеза H_0 не відхиляється.

Побудова критерію для перевірки нульової гіпотези. Нехай H_0 – нульова гіпотеза про розподіл випадкової величини зі значеннями в просторі R^n. Для перевірки H_0 необхідно побудувати (визначити, вибрати) критичну множину S – борелеву множину в R^n таку, що при $\xi\in S$ гіпотеза відхиляється, а при , $\xi\notin S$ – не відхиляється. Борелевих множин S у просторі R^n багато, і, ясна річ, не всі вони будуть однаково слушними для перевірки гіпотези H_0. Для кожної гіпотези H_0 слушними будуть свої критерії S. Ось тут і постає важлива задача (це основна задача перевірки статистичних гіпотез): *як будувати слушні критерії для перевірки гіпотези H_0*.

Щоб розв'язати цю задачу, спочатку розглянемо так звані помилки першого й другого родів та ймовірності цих помилок.

Помилки першого та другого родів. Нехай H_0 – нульова гіпотеза; $\xi(\omega)$ – результат експерименту (реалізація вибірки ξ); S – борелева множина в просторі R^n. Будемо перевіряти H_0, користуючись множиною S як критичною: якщо $\xi(\omega)$ потрапляє до S, то гіпотезу H_0 відхиляємо, а якщо $\xi(\omega)$ не потрапляє до S, то гіпотезу H_0 не відхиляємо.

При цьому можливі такі випадки: гіпотеза H_0 справджується або ні, реалізація $\xi(\omega)$ випадкової величини ξ потрапила до S або ні. Докладніше:

1. *Гіпотеза H_0 справджується*. Реалізація $\xi(\omega)$ не потрапила до S, тому *згідно з критерієм* S *гіпотеза H_0 не відхиляється*.
2. *Гіпотеза H_0 не справджується*. Реалізація $\xi(\omega)$ не потрапила до S; отже, *згідно з критерієм* S *гіпотеза H_0 не відхиляється*.
3. *Гіпотеза H_0 справджується*. Реалізація $\xi(\omega)$ потрапила до S, і *згідно з критерієм* S *гіпотеза H_0 відхиляється*.

4. *Гіпотеза H_0 не справджується*. Реалізація $\xi(\omega)$ потрапила до S, тому *згідно з критерієм* S *гіпотеза H_0 відхиляється*.

З чотирьох випадків два: 1 та 4 – задовільні, а решта (2 і 3) – незадовільні. У випадках 2 та 3 говоримо, що ми припускаємося помилок. При цьому помилки у випадку 2 (гіпотеза H_0 не відхиляється, коли вона не справджується) та у випадку 3 (гіпотеза H_0 відхиляється, коли вона справджується) істотно різняться (за конкретних ситуацій ці помилки, як правило, різняться своєю ціною) і мають свої назви.

Проілюструємо відмінність між названими помилками на прикладі перевірки медичного препарату на токсичність біологічними методами.

Досліджуючи препарат на токсичність, певну його дозу вводять піддослідним тваринам (мишам, кроликам) і реєструють кількість згубних наслідків (зазначимо, що ця кількість є випадковою величиною). Необхідно за реалізацією $\xi(\omega)$ випадкової величини ξ (за кількістю згубних наслідків) зробити висновок про токсичність препарату.

Задачу дослідження препарату на токсичність біологічними методами можна сформулювати в термінах перевірки статистичних гіпотез. А саме, щодо токсичності препарату висуваються гіпотези: H_0 – препарат токсичний та H_1 – препарат нетоксичний. Необхідно перевірити гіпотезу H_0 – відхиляти її або не відхиляти. Вибір між цими діями здійснюється за реалізацією $\xi(\omega)$ випадкової величини ξ – кількістю загиблих тварин.

Як і в кожній задачі перевірки статистичних гіпотез, у розглядуваній можливі помилки:

гіпотеза H_0 не справджується, але згідно з критерієм вона не відхиляється;

гіпотеза H_0 справджується, але згідно з критерієм вона відхиляється.

Подивимося, які наслідки (яка ціна) цих помилок у даній конкретній ситуації.

Гіпотеза H_0 не справджується, але згідно з критерієм не відхиляється. Твердження “гіпотеза H_0 не справджується” у розглядуваній задачі означає, що препарат нетоксичний (не є небезпечним для здоров’я пацієнтів), а твердження “H_0 не відхиляється” означає, що препарат класифікується як токсичний. Таким чином, нетоксичний препарат класифікується як токсичний і повертається постачальнику (для перероблення або знищення). Наслідки помилки такого роду – збитки, зростання вартості товару (ціна помилки – фінансові збитки).

Гіпотеза H_0 справджується, але згідно з критерієм відхиляється. Твердження “H_0 справджується” означає, що препарат токсичний (небезпечний для здоров’я пацієнтів). Твердження “H_0 відхиляється” означає, що препарат класифікується як нетоксичний (не є небезпечним для здоров’я пацієнтів). Таким чином, токсичний препарат (небезпечний для здоров’я пацієнтів) класифікується як нетоксичний (не є небезпечним для здоров’я пацієнтів) і йде на продаж. Наслідком помилки такого роду може стати смерть пацієнта, що вживає цей препарат (ціна помилки – летальний кінець для пацієнта).

Цей приклад наочно показує, що описані вище помилки істотно різняться за своєю ціною.

Означення. Помилка, яка полягає в тому, ідо гіпотеза H_0 відхиляється, коли вона справджується, називається *помилкою першого роду*.

Означення. Помилка, яка полягає в тому, що гіпотеза H_0 не відхиляється, коли вона не справджується, називається *помилкою другого роду*.

Про вибір нульової гіпотези. Раніше вже відзначалося, що нульову гіпотезу із сукупності всіх гіпотез ми вибираємо самі. (*Помилки, пов'язані з перевіркою гіпотез, які полягають у тому, що гіпотезу* H_θ, $\theta \in \Theta$, *ми відхиляємо, коли вона справджується, можна розглядати і не вибравши нульової гіпотези*.) Отже, *за основну (нульову) гіпотезу ми вибираємо ту, для якої важливіше уникнути помилки, що полягає у відхиленні цієї гіпотези, коли вона справджується*. Помилка першого роду – це помилка, якої важливіше уникнути, помилка, ціна якої вища.

Рівень значущості критерію, функція потужності критерію. Чи можна побудувати критерій (критичну множину) для перевірки гіпотези, який би не призводив до помилок? Ні, не можна. Адже, якою б не була критична множина $S \neq \varnothing$, значення $\xi(\omega)$ випадкової величини ξ (результат експерименту) може потрапити до S, Коли гіпотеза H_0 справджується (помилка першого роду), і до $\overline{S}$, коли вона не справджується (помилка другого роду). І оскільки побудувати критерій (критичну множину) для перевірки гіпотези H_0, який би не призводив до помилок, неможливо в принципі, природно, що ми будемо намагатися будувати такі критерії, які мінімізують частоти помилок. Подивимось, як це робиться.

Нехай H_0 – основна гіпотеза. Припустимо для наочності, що вона проста. Конкуруюча гіпотеза тільки одна (позначимо її H_1) і вона також проста. Нехай S – критична множина. Використовуючи S для перевірки гіпотези H_0 (відхиляємо H_0, якщо $\xi \in S$, і не відхиляємо H_0, якщо $\xi \notin S$), ми можемо припуститися помилок двох типів: H_0 – справджується, $\xi \in S$ і, отже, H_0 відхиляється (помилка першого роду); H_0 – не справджується, але $\xi \notin S$ і, отже, H_0 не відхиляється (помилка другого роду). Імовірність помилки першого роду дорівнює ймовірності того, що вибіркове значення ξ потрапить до критичної множини S, коли гіпотеза H_0 справджується, тобто $P\{\xi \in S|H_0\}$ (або, що те саме, $P\{\varphi(\xi)=1|H_0\}$); стисло будемо писати $P(S|H_0)$.

Імовірність помилки другого роду дорівнює ймовірності того, що вибіркове значення потрапить до множини $\overline{S}$, коли гіпотеза H_1 справджується, тобто $P\{\xi \in \overline{S}|H_1\}$ (або, що те саме, $P\{\varphi(\xi)=0|H_1\}$); стисло будемо писати $P(\overline{S}|H_1)$. Зазначимо, що $P\{\xi \in S|H_0\}$ – це ймовірність події $\{\xi \in S\}$, обчислена в припущенні, що гіпотеза H_0 справджується; вона не має нічого спільного з умовною ймовірністю (на H_0 та H_1 і зручно дивитися як на індекси P).

Імовірності помилок першого та другого родів однозначно визначаються критичною множиною S (а отже, й частоти помилок першого і другого родів однозначно визначаються цією множиною). Тому, вибираючи S, скажімо, за умови, що ймовірність помилки першого роду мала, ми водночас одержимо ймовірність помилки другого роду, яка визначається вибраною критичною множиною S і буде такою, якою вийде. Можна вибрати S також за умови, що ймовірність помилки другого роду мала, але при цьому ймовірність помилки першого роду визначиться вибраною множиною S і буде такою, якою вийде. Отже, вибрати S так, щоб одночасно були контрольовані ймовірності помилок першого та другого родів (а разом із ними і частоти помилок першого та другого родів), не вдасться. Тому будемо діяти так. Оскільки важливіше уникнути помилки першого роду (її ціна вища), першою вимогою до критерію φ (до критичної множини S) буде така: ймовірність помилки першого роду $P(S|H_0)$ має бути малою. Це означає, що, використовуючи критерій S у довгій серії експериментів, гіпотезу H_0, коли вона справджується, будемо відхиляти зрідка.

Таким чином, перша вимога щодо критерію – фіксуємо мале α і вибираємо критерій φ (критичну множину S) так, щоб імовірність помилки першого роду не перевищувала α:

$$P(S|H_0)\le\alpha \ (P\{\varphi(\xi)=1|H_0\}\le\alpha). \qquad (6)$$

Якщо ця основна вимога задовольняється більш як одним способом, то остаточний вибір критерію здійснюється так, щоб імовірність $P(S|H_1)$відхилення гіпотези H_0, коли вона не справджується, була максимальною (а отже, частота того, що гіпотеза H_0 відхиляється, коли вона не справджується, була максимальною), для чого множину S вибираємо якомога ширшою.

Означення. Число α, що обмежує зверху ймовірність помилки першого роду, називається рівнем значущості.

Якщо критична множина S (критерій φ) задовольняє умову

$$P(S|H_0)\le\alpha \ (P\{\varphi(\xi)=1|H_0\}\le\alpha), \qquad (7)$$

то будемо говорити, що критична множина S (критерій φ) відповідає рівню значущості α.

Означення. Ймовірність $P(S|H_1)$ відхилити основну гіпотезу, коли справджується альтернативна гіпотеза H_1, будемо називати потужністю критерію S.

Якщо альтернативна гіпотеза складна (причому коли вона справджується, справджується одна з простих гіпотез H_θ, $\theta\in\Theta$), то для кожного $\theta\in\Theta$ можна обчислити $P(S|H_\theta)$.

Означення. Функція

$$\beta(\theta) = P(S|H_\theta), \qquad (8)$$

яка для кожного $\theta \in \Theta$ дорівнює ймовірності відхилити основну гіпотезу H_0, коли справджується гіпотеза H_θ, називається *функцією потужності критерію*.

Зауваження. Питання про те, яким має бути рівень значущості, не є статистичною задачею. Здебільшого за рівень значущості приймають числа 0,1; 0,05; 0,01; 0,001. Чим серйозніші наслідки помилки першого роду, тим меншим має бути рівень значущості.

Приклад 1.1. *Партія кількістю* N=10000 *виробів, призначена для продажу, проходить вибірковий контроль на якість. Постачальник упевнений, що частка дефектних виробів становить* 1% *(або менше), і бажає, щоб кожного разу, коли це дійсно так, імовірність того, що партія витримує контроль, дорівнювала* 0,9. *Покупець вважає, що партію доцільно закупити навіть тоді, коли частка дефектних виробів буде перевищувати 1%, але* 6% *дефектних виробів він вважає гранично допустимою часткою і бажає, щоб контроль виявляв це з імовірністю* 0,95. *Постачальник і покупець домовилися контроль здійснювати так: із партії* 10000 *виробів беруть випадкову вибірку обсягом n. Якщо при цьому частка дефектних виробів* $\xi(\omega)$ *мала* ($\xi(\omega) \le l$), *то партія витримує контроль і закуповується, у противному разі – ні.*

Якими мають бути обсяг вибірки n партії та значення l?

Відповісти на це запитання, сформулювавши і розв'язавши поставлену задачу як задачу перевірки статистичних гіпотез.

Розв'язання. Робити висновок про закупівлю або відхилення партії виробів будемо за реалізацією $\xi(\omega)$ випадкової величини ξ – кількістю дефектних виробів із n вибраних. Оскільки обсяг партії великий (N = 10000), можна вважати події, пов'язані з відбиранням виробів, на кожному кроці незалежними, а отже, можемо вважати, що випадкова величина ξ має біномний розподіл

$$P\{\xi=k\} = C_n^k p^k (1-p)^{n-k},\ k= 0,1,...,n, \qquad (9)$$

де p – ймовірність вибору дефектного виробу. Нас цікавлять значення параметру p: p_1 = 0,01 (відповідає 1% дефектних виробів у партії) та p_2 = 0,06 (відповідає 6% дефектних виробів у партії). Інакше кажучи, становлять інтерес такі розподіли випадкової величини ξ:

$$P_1(k) = C_n^k p_1^k (1-p_1)^{n-k},\ k= 0,1,...,n;\ \ p_1 = 0{,}01; \qquad (10)$$

$$P_2(k) = C_n^k p_2^k (1-p_2)^{n-k},\ k= 0,1,...,n;\ \ p_2 = 0{,}06. \qquad (11)$$

Тим самим щодо розподілу випадкової величини ξ ми висунули дві гіпотези: H_1 – розподілом ξ є P_1 та H_2 – розподілом ξ є P_2.

Необхідно за кількістю дефектних виробів $\xi(\omega)$ (за реалізацією $\xi(\omega)$ біномно розподіленої випадкової величини ξ) дійти висновку: який із розподілів, P_1 чи P_2, є розподілом ξ.

Щоб розв'язати цю задачу, відповідно до викладеної методики із сукупності можливих гіпотез $\wp=\{H_1,H_2\}$ вибираємо основну (нульову), а потім перевіряємо її. Для того, щоб вибрати нульову гіпотезу з $\{H_1,H_2\}$ подивимось, якими будуть наслідки помилок при відхиленні гіпотез H_1 і H_2, коли вони справджуються.

Відхилення гіпотези H_1, коли вона справджується, означає класифікацію партії, що містить 1% дефектних виробів, як такої, що містить 6%; така помилка небажана для постачальника. Відхилення гіпотези H_2, коли вона справджується, означає класифікацію партії, що містить 6% дефектних виробів, як такої, що містить 1%; така помилка небажана для покупця. Тому в розглядуваній задачі не має значення, яку з гіпотез, H_1 чи H_2, вибрати за основну, якщо тільки ми не є зацікавленою стороною. Для постачальника доцільно віддати перевагу – як основній – гіпотезі H_1, для покупця – гіпотезі H_2.

Розглянемо – як основну – гіпотезу H_2. Отже, H_0 це гіпотеза: кількість дефектних виробів з n вибраних має біномний розподіл з параметром 0,06, тобто

$$\mathrm{P}_0(k) = \mathrm{C}_n^k(0{,}06)^k(0{,}94)^{n-k},\ k= 0,1,...,n; \tag{12}$$

Як критичну множину для перевірки гіпотези H_0 природно розглядати $\mathrm{S}=\{k\colon k \le l\}$. Ця множина цілком визначається вибором числа l.

Зазначимо, що вибір нульової гіпотези і критичної множини фактично визначається домовленістю постачальника та покупця: якщо частка дефектних виробів $\xi(\omega)$ серед n вибраних мала ($\xi(\omega) \le l$), то партія витримує контроль і закуповується, в противному разі – ні.

За умовою задачі партія, що містить 6% дефектних виробів (справджується гіпотеза H_0), має виявлятися ,(класифікуватися як партія, що містить 6% дефектних виробів) з імовірністю 0,95. А отже, з імовірністю 0,05 партія, що містить 6% дефектних виробів (гіпотеза H_0 справджується), має класифікуватися як партія, що містить 1% дефектних виробів (H_0 відхиляється). Останнє означає, що рівень значущості α критерію – число, яке обмежує зверху ймовірність помилки першого роду, – має дорівнювати 0,05, тобто

$$\mathrm{P}(\mathrm{S}|H_0) = \mathrm{P}_0(\xi\in\mathrm{S}) = \mathrm{P}_0(\xi \le l) = \sum_{k=0}^{l}\mathrm{C}_n^k(0{,}06)^k(0{,}94)^{n-k} \le \alpha = 0{,}05. \tag{13}$$

Якщо партія містить 1% дефектних виробів (справджується гіпотеза H_1), то вона витримує контроль (класифікується як партія, що містить 1% дефектних виробів) з імовірністю 0,9, тобто гіпотеза H_0 має відхилятися з імовірністю 0,9; тому

$$P(S|H_1) = P_1(\xi \in S) = P_1(\xi \le l) = \sum_{k=0}^{l} C_n^k (0{,}01)^k (0{,}99)^{n-k} \ge 0{,}9. \quad (14)$$

Таким чином, для того щоб знайти n та l, маємо співвідношення

$$\sum_{k=0}^{l} C_n^k (0{,}06)^k (0{,}94)^{n-k} \le 0{,}05; \qquad (15)$$

$$\sum_{k=0}^{l} C_n^k (0{,}01)^k (0{,}99)^{n-k} \ge 0{,}9. \qquad (16).$$

Електронний документ для обчислень доцільно організувати так, як показано на рис. 1.

СУММ | =БИНОМРАСП(H2;$G$2;$F$2;ЛОЖЬ)

	A	B	C	D	E	F	G	H	I	J	K
1	k	P_k=P{ξ=k}	∑P_k		p1	p2	n	k	P_k=P{ξ=k}	∑P_k	
2	0	0,605006	0,605006		0,01	0,06	50	0	=БИНОМРАСП(H2;$G$2;$F$2;		
3	1	0,305559	0,910565					1	ЛОЖЬ)		
4	2	0,075618	0,986183	∑P_k>=0,9	l1	l2	∑P_k<=0,05	2	0,226243	0,416246	
5	3	0,012221	0,998404	l>=l1	1	0	l<=l2	3	0,231057	0,647303	
6	4	0,00145	0,999854					4	0,173293	0,820596	
7	5	0,000135	0,999989					5	0,101763	0,922359	
8	6	1,02E-05	0,999999					6	0,048717	0,971076	
9	7	6,48E-07	1					7	0,019546	0,990622	
10	8	3,52E-08	1					8	0,006706	0,997328	
11	9	1,66E-09	1					9	0,001997	0,999325	
12	10	6,87E-11	1					10	0,000523	0,999848	
13	11	2,52E-12	1					11	0,000121	0,999969	
14	12	8,29E-14	1					12	2,52E-05	0,999994	
15	13	2,45E-15	1					13	4,7E-06	0,999999	

БИНОМРАСП(число_успехов; число_испытаний; вероятность_успеха; интегральная)

Рис. 1. Організація обчислень для перевірки гіпотез

Обчислення сум у формулах (16), (15) реалізується у стовпцях A-C та I-K відповідно. Розглянемо заповнення стовпців A-C; для стовпців I-K обчислення організовуються аналогічно. В перший стовпець заносяться значення змінної k=0,1,…,n. У другому стовпці обчислюються значення імовірностей $P_1(k)$ за формулою (10). Для цього використовується стандартна функція БИНОМРАСП, яка приймає чотири аргументи: значення k (число успіхів у схемі незалежних випробувань Бернуллі), число випробувань n, імовірність успіху p та логічний показник, що визначає, чи обчислювати всю суму $\sum_{i=0}^{k} P(i)$ (значення параметру «істинне»), чи тільки окремий її доданок $P(k)$ (значення параметру «хибне»). В комірку B2 введемо формулу =БИНОМРАСП(A2;$G$2;$E$2;ЛОЖЬ), де k розташоване в комірці A2, число випробувань n – в комірці G2 (абсолютна адресація), імовірність успіху p_1 – в комірці F2 (також абсолютна адресація). Формулу «протягнемо» на весь другий стовпець. В третьому стовпці обчислимо суму (16). В комірку C2 введемо формулу =СУММ($B$2:B2) і протягнемо її на весь третій стовпець. В принципі, можна було б обійтися тільки другим стовпцем, встановивши останній параметр в функції БИНОМРАСП у значення ИСТИНА.

В комірках I-K обчислення проводяться аналогічно, тільки у функції БИНОМРАСП у другому стовпці в якості другого аргументу використовується (абсолютне) посилання на комірку F2, що містить значення імовірності p_2.

Визначаючи n і l, ми, природно, будемо вибирати n якомога меншим. Процес вибору розміру вибірки n (і відповідного йому значення l) буде таким. Розв‘язком нерівностей (16) і (15) будуть значення $l \geq l_1$ та $l \leq l_2$ (з певними величинами l_1 і l_2) відповідно. Якщо для деякого n буде $l_1 > l_2$, то це означатиме, що для даного розміру вибірки неможливо вибрати жодного l, яке б одночасно задовольняло нерівності (15), (16), тобто було б прийнятним для постачальника і покупця. В цьому випадку розмір вибірки n необхідно збільшити.

Якщо маємо $l_2 > l_1$, то система (15), (16) задовольняється для більш ніж одного значення $l = l_1, \ldots, l_2$, що означає можливість зменшення розміру вибірки n.

Тільки у випадку $l_1 = l_2$, значення n буде оптимальним, і пара n, $l = l_1 = l_2$ і буде розв‘язком задачі пошуку найкращого критерію перевірки гіпотез.

Значення l_1 вибираємо як найменша величина k, для якої значення у стовпці C перевищують 0,9. Величину l_2 вибираємо як найбільше значення k, для якого значення у стовпці I не перевищують 0,05.

Подивимося, чи можна при n=50 вибрати l так, щоб задовольнялися нерівності (15), (16). Зі стовпчика C бачимо, що $\sum_{k=0}^{1} P_1(k) = 0{,}9106 \geq 0{,}9$, тобто $l_1 = 1$. Аналізуючи дані стовпця J, отримаємо $\sum_{k=0}^{0} P_2(k) = 0{,}0453 \leq 0{,}05$, тобто $l_2 = 0$. Отже, $l_1 > l_2$, і розв’язку задачі не існує; розмір вибірки n=50 виявляється замалим.

Спробуємо збільшити значення n, наприклад, n=100. Маємо: $\sum_{k=0}^{2} P_1(k) = 0{,}9206 \geq 0{,}9$, тобто $l_1 = 2$; $\sum_{k=0}^{1} P_2(k) = 0{,}0152 \leq 0{,}05$, тобто $l_2 = 1$. Знову $l_1 > l_2$, і n=100 також замало.

Візьмемо n=150. Для цього значення $\sum_{k=0}^{3} P_1(k) = 0{,}9353 \geq 0{,}9$, тобто $l_1 = 3$; $\sum_{k=0}^{4} P_2(k) = 0{,}0499 \leq 0{,}05$, тобто $l_2 = 4$. В цьому випадку в якості l можливий вибір між значеннями 3, 4.

Спробуємо трохи зменшити значення n. Нехай n=130. Маємо: $\sum_{k=0}^{3} P_1(k)$ $=0{,}9578 \geq 0{,}9$, тобто $l_1=3$; $\sum_{k=0}^{3} P_2(k)=0{,}0438 \leq 0{,}05$, тобто $l_2=3$. Оскільки $l_1=l_2=3$, то пара n=130, l=3, і є шуканим розв'язком задачі, задовольняючи одночасно нерівності (14), (15), тобто вимоги постачальника і покупця. Критичною множиною (критерієм) буде

$$S = \{x\colon 0 \leq x \leq 3\}. \tag{17}$$

Для цього критерію ймовірність помилки першого роду

$$P(S|H_0) = P_0(\xi \in S) = P_0(\xi \leq 3) = \sum_{k=0}^{3} C_n^k (0{,}06)^k (0{,}94)^{n-k} = 0{,}0438, \tag{18}$$

а ймовірність помилки другого роду

$$P(\overline{S}\,|H_1) = P_1(\xi \in \overline{S}\,) = P_1(\xi > 3) = \sum_{k=4}^{n} C_n^k (0{,}01)^k (0{,}99)^{n-k} =$$

$$= 1 - \sum_{k=0}^{3} C_n^k (0{,}01)^k (0{,}99)^{n-k} = 1 - 0{,}9578 = 0{,}0422. \tag{19}$$

Потужність критерію

$$P(S|H_1) = 1 - P(\overline{S}\,|H_1) = 1 - 0{,}0422 = 0{,}9578. \tag{20}$$

Імовірність помилки першого роду, що дорівнює 0,0438, можна інтерпретувати так. При використанні описаного критерію для партій виробів, які містять 6% браку, в середньому 4 партії зі 100 будуть класифікуватися як такі, що містять 1% браку.

Імовірність помилки другого роду, що дорівнює 0,0422, можна інтерпретувати так. Серед партій, які містять 1% браку, в середньому 4 партії зі 100 будуть класифікуватися як такі, що містять 6% браку.

Значення потужності критерію $P(S|H_1)$=0,9578 інтерпретується так. Серед партій, які містять 1% браку, в середньому 96 партій зі 100 будуть класифікуватися як такі, що містять 1% браку.

Приклад 1.2. *Попередніми дослідженнями встановлено, що інтенсивність транспортного потоку на автомобільній магістралі складає* λ_0*=100 авто/хв. Для перевірки цього було встановлено спостереження за допомогою відеокамери-фіксатора проїжджаючих авто. За першу хвилину спостережень перед камерою проїхало 79 авто. Чи суперечать отримані дані результатам попередніх досліджень, тобто інтенсивність транспортного* *потоку дійсно складає $\lambda=\lambda_0=100$ авто/хв.? Задачу розв'язати за допомогою перевірки статистичних гіпотез. Як зміниться результат, якщо відомо, що за другу хвилину спостережень перед камерою проїхало 120 авто? Побудувати критерії перевірки гіпотез для випадку одного та двох хвилин спостережень. Прийняти рівень значущості критеріїв α=5%. Який з двох критеріїв буде потужнішим і чому?*

Припустимо, що автомобільний потік моделюється простішим пуассонівським потоком, для якого кількість подій ξ_1 (проїздів автомобілів) в будь-який заданий інтервал часу (в даному випадку – одну хвилину) характеризується ймовірностями

$$P(k) = P\{\xi_1=k\} = \frac{\lambda^k}{k!}e^{-\lambda}, \quad k=0,1,2,\ldots,n,\ldots \tag{21}$$

де λ – інтенсивність потоку, тобто середнє значення кількості подій в одиницю часу. Цей параметр в умовах нашої задачі і є невідомим; для його оцінки застосуємо метод перевірки статистичних гіпотез.

Природно щодо цього параметру сформулювати основну (просту) гіпотезу H_0: $\lambda=\lambda_0$; в якості альтернативної використати складну гіпотезу H_1: $\lambda\neq\lambda_0$, яку можна вважати параметричною, що залежить від параметру λ і позначатиметься як H_λ.

Для побудови критерію перевірки зауважимо, що спостережуване значення ξ_1 для потоку інтенсивності λ не повинно значно відрізнятись від числа λ. Формально це можна записати у такий спосіб: якщо значення ξ_1 належить множині $\overline{S}=\{\lambda_0-l,\ldots,\lambda_0+l\}$, гіпотезу H_0 ми не відхиляємо, в противному випадку – відхиляємо. Критичною для такого критерію буде множина $S = Z_0 \setminus \overline{S} = \{0,1,\ldots, \lambda_0-l-1\}\cup\{\lambda_0+l+1,\ldots\}$, де Z_0 – множина цілих невід'ємних чисел. Індикативна функція цього критерію має вигляд

$$I_S(x) = \varphi(x) = \begin{cases} 1, \text{ якщо } 0 \le x \le \lambda_0-l-1 \text{ або } x \ge \lambda_0+l+1, \\ 0, \text{ якщо } \lambda_0-l \le x \le \lambda_0+l. \end{cases} \tag{22}$$

Критерій залежить від параметру l. Помилка першого роду під час застосування цього критерію полягає в тому, що спостережена кількість авто або значно менша, ніж дійсне значення λ_0, або ж значно його перевищує, тобто $\xi_1\in S$. Ймовірність цієї похибки повинна не перевищувати заданої величини α – рівня значущості критерію

$$P(S|H_0) = P(\xi_1\in S) = P\{0 \le \xi_1 \le \lambda_0-l-1\} + P\{\xi_1 \ge \lambda_0+l+1\} =$$

$$= 1 - P(\overline{S}|H_0) = 1 - P\{\lambda_0-l \le x \le \lambda_0+l\} = 1-\sum_{k=\lambda_0-l}^{\lambda_0+l}\frac{\lambda_0^k}{k!}e^{-\lambda_0} \le \alpha \tag{23}$$

або, що еквівалентно,

$$P(\overline{S}\,|H_0) = \sum_{k=\lambda_0-l}^{\lambda_0+l} \frac{\lambda_0^k}{k!} e^{-\lambda_0} \geq 1 - \alpha. \qquad (24)$$

Розв'язком цієї нерівності буде множина значень $l_1 \leq l \leq \lambda_0$ з певним нижнім граничним значенням l_1. Критерії з такими параметрами l будуть мати заданий рівень значущості α. Який з цих критеріїв треба обрати як оптимальний? Очевидно, це буде той з них, що мінімізує ймовірність $P(\overline{S}\,|H_1)$ похибки другого роду, або, що теж саме, максимізує по параметру l функцію потужності критерію

$$\beta_1(\lambda;l) = P(S|H_\lambda) = 1 - P(\overline{S}\,|H_\lambda) = 1 - \sum_{k=\lambda-l}^{\lambda+l} \frac{\lambda^k}{k!} e^{-\lambda} \to \max \text{ (по } l\text{)}. \qquad (25)$$

Максимізує цю функцію найменше з можливих значення l, тобто значення l_1, оскільки для будь-якого $l > l_1$

$$\beta_1(\lambda;l_1) - \beta_1(\lambda;l) = \sum_{k=\lambda-l+1}^{\lambda-l_1} \frac{\lambda^k}{k!} e^{-\lambda} + \sum_{k=\lambda+l_1+1}^{\lambda+l} \frac{\lambda^k}{k!} e^{-\lambda} > 0. \qquad (26)$$

Отже, гіпотезу H_0: $\lambda=\lambda_0=100$ ми не відхиляємо, якщо $\lambda_0-l_1 \leq \xi_1 \leq \lambda_0+l_1$ (l_1 визначається з нерівності (24)) і відхиляємо в противному випадку.

Розглянемо тепер випадок, коли рішення про відхилення або не відхилення гіпотези H_0 формується на базі двох спостережень: за першу хвилину проїхало ξ_1 авто, а за другу – ξ_2 авто. Оскільки транспортний потік вважається простішим, то ці випадкові величини є незалежними і однаково розподіленими за законом (21).

Критерій перевірки гіпотези H_0 побудуємо у такий спосіб. Будемо не відхиляти гіпотезу H_0, якщо ξ_1 і ξ_2 не сильно відрізняються від значення λ_0. Тобто гіпотеза не відхиляється, тільки якщо одночасно $\xi_1 \in [\lambda_0-l;\ \lambda_0+l]$ та $\xi_2 \in [\lambda_0-l;\ \lambda_0+l]$, а якщо, скажімо $\xi_1 \in [\lambda_0-l;\ \lambda_0+l]$, а $\xi_2 \notin [\lambda_0-l;\ \lambda_0+l]$, або $\xi_2 \in [\lambda_0-l;\ \lambda_0+l]$, а $\xi_1 \notin [\lambda_0-l;\ \lambda_0+l]$ – гіпотеза відхиляється. Критична множина S та множина $\overline{S}$ в даному випадку будуть двовимірними, тобто складаються з упорядкованих пар (x_1,x_2), де x_1,x_2 – невід'ємні цілі числа. Множина $\overline{S} = \{\lambda_0-l,\ldots,\lambda_0+l\} \otimes \{\lambda_0-l,\ldots,\lambda_0+l\}$; символ $\otimes$ означає декартовий добуток двох множин, тобто множину упоряднених пар, в яких перший елемент належить першій множині, а другий елемент – другій. Тоді критична множина (випадок, коли ми відхиляємо нульову гіпотезу) визначиться як $S=Z_0^2\backslash\overline{S}$, де $Z_0^2=Z_0\otimes Z_0$. Ймовірність похибки першого роду, яка не повинна перевищувати заданого рівня значущості α, обчислимо з урахуванням незалежності випадкових величин ξ_1 і ξ_2:

$$P(S|H_0) = P\{(\xi_1,\xi_2)\in S\} = 1 - P\{(\xi_1,\xi_2)\in \overline{S}\} = 1 - P(\overline{S}|H_0) =$$

$$= 1 - P\{\lambda_0 - l \le \xi_1 \le \lambda_0 + l\}\cdot P\{\lambda_0 - l \le \xi_2 \le \lambda_0 + l\} = 1 - \left(\sum_{k=\lambda_0-l}^{\lambda_0+l}\frac{\lambda_0{}^k}{k!}e^{-\lambda_0}\right)^2 \le \alpha. \quad (27)$$

Таким чином, для визначення l маємо нерівність

$$P(\overline{S}|H_0) = \left(\sum_{k=\lambda_0-l}^{\lambda_0+l}\frac{\lambda_0{}^k}{k!}e^{-\lambda_0}\right)^2 \ge 1 - \alpha. \quad (28)$$

Розв’язком цієї нерівності буде множина значень $l_2 \le l \le \lambda_0$ з певним нижнім граничним значенням l_2. Критерії з такими параметрами l будуть мати заданий рівень значущості α. Як і в попередньому випадку, оптимальним (має найбільшу потужність) буде критерій із значенням $l=l_2$. Як будуть співвідноситись між собою значення параметрів критеріїв l_1 та l_2? Ймовірність $P(\overline{S}|H_0)$ для обох критеріїв виражається через суму у правій частині нерівності (24), тільки у формулі (28) ця сума стоїть у квадраті. Оскільки ця сума менше одиниці, то її квадрат буде ще меншим, і щоб їх урівняти, необхідно взяти більше доданків у сумі (28) порівняно із сумою (24). Отже, ми маємо $l_1 \le l_2$. Ця ситуація є досить типовою в статистиці: залучення нової інформації про досліджувану величину дозволяє дещо послабити обмеження на множину $\overline{S}$ не відхилення нульової гіпотези.

Функція потужності критерію (ймовірність відхилити основну гіпотезу, коли справджується альтернативна) буде виражатись формулою, аналогічною (25)

$$\beta_2(\lambda;l) = P(S|H_\lambda) = 1 - P(\overline{S}|H_\lambda) = 1 - \left(\sum_{k=\lambda-l}^{\lambda+l}\frac{\lambda^k}{k!}e^{-\lambda}\right)^2. \quad (29)$$

Варіант можливої організації електронного документу для обчислень показано на рис. 2.

ADDRESS =1-SUM((INDEX($B$2:$B$202;$E$2+1-I5):INDEX($B$2:$B$202;$E$2+1+I5)))^2

	A	B	E	F	I	J	M	N	P	Q	S	T
1	k	Pk	λ0				k	Pk	λ			
2	0	3,72008E-44	100				0	1,4789E-74	170			
3	1	3,72008E-42	Рівень значущості				1	2,51413E-72	Потужність критерію			
4	2	1,86004E-40	l	P{S|H0}	l	P{S|H0}	2	2,13701E-70	l	P{S|H0}	l	P{S|H0}
5	3	6,20013E-39	0	0,960139003	0	=1-SUM((INDEX($B$2:$B$202;$E$2+1-I5):INDEX($B$2:$B$202;$E$2+1+I5)))^2				1	0	1
6	4	1,55003E-37	1	0,880811673	1	0,985794	4	5,14662E-67	1	1	1	1
7	5	3,10006E-36	2	0,802656802	2	0,961056	5	1,74985E-65	2	1	2	1
8	6	5,16677E-35	3	0,726418145	3	0,925153	6	4,95792E-64	3	1	3	1
9	7	7,3811E-34	4	0,652784509	4	0,879441	7	1,20407E-62	4	1	4	1
10	8	9,22638E-33	5	0,582371424	5	0,825586	8	2,55864E-61	5	1	5	1
11	9	1,02515E-31	6	0,515706168	6	0,765459	9	4,83298E-60	6	1	6	1
12	10	1,02515E-30	7	0,453216743	7	0,701028	10	8,21607E-59	7	1	7	1
13	11	9,31957E-30	8	0,395225145	8	0,634247	11	1,26976E-57	8	1	8	1
14	12	7,76631E-29	9	0,341945017	9	0,566964	12	1,79882E-56	9	1	9	1
15	13	5,97409E-28	10	0,293483523	10	0,500834	13	2,35231E-55	10	0,999999	10	1
16	14	4,2672E-27	11	0,249847077	11	0,437271	14	2,85637E-54	11	0,999999	11	1
17	15	2,8448E-26	12	0,21095035	12	0,377401	15	3,23722E-53	12	0,999999	12	1
18	16	1,778E-25	13	0,176627862	13	0,322058	16	3,43955E-52	13	0,999998	13	1
19	17	1,04588E-24	14	0,146647368	14	0,271789	17	3,43955E-51	14	0,999997	14	1
20	18	5,81046E-24	15	0,120724247	15	0,226874	18	3,24846E-50	15	0,999995	15	1
21	19	3,05814E-23	16	0,098536104	16	0,187363	19	2,90652E-49	16	0,999993	16	1
22	20	1,52907E-22	17	0,079736889	17	0,153116	20	2,47054E-48	17	0,999989	17	1
23	21	7,28128E-22	18	0,063969952	18	0,123848	21	1,99996E-47	18	0,999985	18	1
24	22	3,30967E-21	19	0,050879571	19	0,09917	22	1,54542E-46	19	0,999977	19	1
25	23	1,43899E-20	20	0,040120652	20	0,078632	23	1,14227E-45	20	0,999968	20	1
26	24	5,99579E-20	21	0,031366447	21	0,061749	24	8,09108E-45	21	0,999954	21	1
27	25	2,39831E-19	22	0,024314254	22	0,048037	25	5,50193E-44	22	0,999934	22	1
28	26	9,22429E-19	23	0,018689198	23	0,037029	26	3,59742E-43	23	0,999907	23	1
29	27	3,4164E-18	24	0,014246263	24	0,02829	27	2,26504E-42	24	0,99987	24	1

Рис. 2. Організація обчислень на листі електронної таблиці.

Значення P(k) за формулою (21) розташуємо у стовпці B, увівши в комірку B2 формулу =POISSON.DIST(A2;$E$2;FALSE) (у російськомовній версії ПУАССОН.РАСП) та протягнувши її на весь стовпець (на 200 значень). Функція POISSON.DIST повертає значення ймовірності P(k) розподілу Пуассона; її перший аргумент – значення k, другий – значення λ, третій (логічний) задає, чи обчислювати суму ймовірностей від нуля до k (значення TRUE), або тільки окреме значення P(k) (значення FALSE).

Значення похибки $P(S|H_0)$ першого роду для критерію, де в якості даних використовується тільки значення ξ_1 (далі буде позначатись як критерій I), обчислимо у стовпці F за формулою (23). Для формування діапазону комірок, в який обчислюється сума $\sum_{k=\lambda-l}^{\lambda+l}\frac{\lambda^k}{k!}e^{-\lambda}$, зручно застосувати функцію INDEX (у російськомовній версії ИНДЕКС), яка повертає посилання на комірку, розташовану у певному рядку (другий параметр) заданого діапазону (перший параметр). Отже, в комірку F5 введемо формулу 1-SUM((INDEX($B$2:$B$202;$E$2+1-E5):INDEX($B$2:$B$202;$E$2+1+E5))) і протягнемо її на весь стовпець F.

Аналогічні обчислення похибки $P(S|H_0)$ першого роду для критерію, де в якості даних використовуються ξ_1 і ξ_2 (в подальшому позначатиметься як критерій II), виконуються у стовпці J згідно з формулою (27) (суму необхідно піднести у квадрат).

Знайдемо найменше l у стовпці E, для якого відповідне значення ймовірності $P(S|H_0)$ буде менше заданого рівня значущості 0,05; таке значення l_1=20, для якого

$$P(S|H_0)=0{,}04012. \qquad (30)$$

Отже, критерій I сформулюється так: якщо спостережуване значення кількості автомобілів знаходиться в межах від 80 до 120, гіпотезу H_0: λ=100 ми не відхиляємо, в противному випадку – відхиляємо. Такий критерій у випадку, якщо інтенсивність потоку дійсно дорівнює 100 авто/хв., буде помилятись приблизно в 4 випадках зі 100 (див. формулу (30)).

В нашому випадку ξ_1 = 79, то згідно з критерієм гіпотезу H_0 нам треба відхилити.

Аналогічно для критерію II у стовпці I знайдемо значення l, для якого значення P(S|H_0) у стовпці J менше 0,05. Це буде значення l_2=22, для якого

$$P(S|H_0)= 0{,}048. \qquad (31)$$

Таким чином, критерій II може бути сформульований у такий спосіб: якщо значення ξ_1 і ξ_2 одночасно не менше 78 і не більше 122, то гіпотезу H_0: λ=100 не відхиляємо, у противному випадку – відхиляємо. Такий критерій у випадку, якщо інтенсивність потоку дійсно дорівнює 100 авто/хв., буде помилятись приблизно в 5 випадках зі 100 (див. формулу (31)). В нашому випадку ξ_1 = 79, ξ_2 = 120, то згідно з критерієм гіпотезу H_0 нам треба не відхилити. Як бачимо, додаткова інформація про спостережувану величину здатна кардинально змінити оцінку досліджуваного параметру.

Який з критеріїв (I або II) є кращим? Звичайно, при заданому рівні значущості (ймовірності помилки першого роду) кращим буде той, що менше робить помилок другого роду; такою характеристикою якості є потужність критерію (ймовірність відхилити основну гіпотезу, коли справджується альтернативна гіпотеза H_λ). Для критеріїв I і II відповідні потужності критеріїв виражаються формулами (25), (29).

Обчислення, необхідні для отримання значень потужності критеріїв розмістимо у правій частині листа електронної таблиці (див. рис. 2). У стовпці N розмістимо значення ймовірностей розподілу Пуассона P(k) (аналогічно стовпцю B) для заданого значення λ, яке розташуємо у комірці P2. В комірку N2 введемо формулу =POISSON.DIST(M2;$P$2;FALSE) (у російськомовній версії ПУАССОН.РАСП) і протягнемо її на весь стовпець.

Формули для обчислення потужностей критеріїв I та II за формулами (25), (29) розмістимо у стовпцях Q і T відповідно. В комірку Q5 введемо формулу =1-SUM((INDEX($N$2:$N$202;$E$2+1-P5):INDEX($N$2:$N$202;$E$2+1+P5))) та протягнемо її далі на стовпець; в комірку T5 введемо формулу =1-SUM((INDEX($N$2:$N$202;$E$2+1-S5):INDEX($N$2:$N$202;$E$2+1+S5)))^2 і також протягнемо її. Відповідні значення потужностей для критеріїв I та II побачимо у комірках, що відповідають значенням l=20 та l=22 відповідно.

Проведемо обчислення для різних значень параметру λ в комірці P2. Скажімо, переберемо всі значення, починаючи з λ=30 до λ=170 з кроком $\Delta\lambda$=10. Результати зведемо в таблицю.

Таблиця 1. Потужності критеріїв I та II як функції параметру λ

λ	Потужність критерію β_1	Потужність критерію β_2
30	1	1
40	1	1
50	0,999943	1
60	0,992182	0,999788
70	0,870901	0,96619
80	0,485143	0,635884
90	0,134244	0,175505
100	0,040121	0,048037
110	0,159466	0,222777
120	0,47579	0,644989
130	0,796409	0,933432
140	0,952961	0,995487
150	0,993445	0,999888
160	0,999429	0,999999
170	0,999968	1

Відповідна діаграма показана на рис. 3.

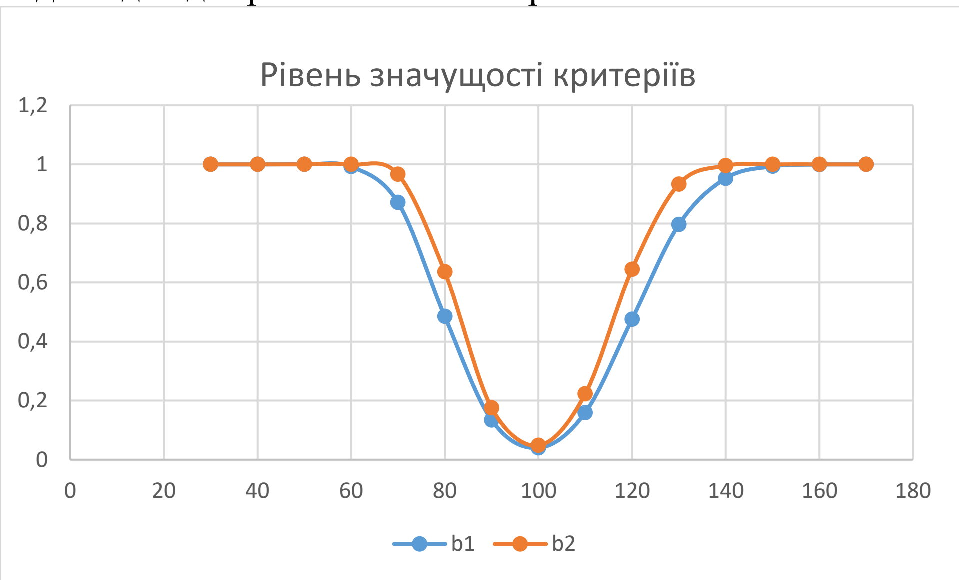

Рис. 2. Графіки потужності критеріїв I і II в залежності від параметру λ.

З графіка видно, що критерії добре «відрізняють» значення λ, які менше 70 або більші 130 від заданого значення λ_0=100. Але графік потужності для критерію II всюди лежить вище графіку для критерію I. Наприклад, критерій II відхиляє основну гіпотезу H_0: λ=100 у випадку, коли дійсне значення λ=70, приблизно у 97 випадках із 100, в той час як для критерію I цей показник складає 87 випадків зі 100.

Отже, критерій II виявляється більш якісним, ніж критерій I, і це природно, тому що він залучає більшу інформацію для аналізу, ніж критерій I.

1.2. Задачі

1. Щоб з'ясувати, чи є хвороба заразною, біолог прищеплює її п'яти мишам і розміщує їх в одній клітці із те нещепленими мишами. На його думку, якщо хвороба перехідна, то ймовірність того, що миша захворіє протягом 10 днів, становить 0,9; якщо ж хвороба неперехідна, то ймовірність того, що миша захворіє протягом зазначеного часу, становить 0,05. Через 10 днів фіксується кількість нещеплених мишей з ознаками хвороби.

Яку мінімальну кількість нещеплених мишей необхідно використати в експерименті, щоб з імовірністю, не ; меншою ніж 0,999, виявити перехідну хворобу і з імовірністю, не більшою ніж 0,01, класифікувати неперехідну хворобу як перехідну.

Сформулювати та розв'язати поставлену задачу як задачу перевірки статистичних гіпотез (див. задачу 8).

2. Маємо монету, яку треба дослідити на симетричність (вона може бути як симетричною – імовірність випадання «герба» становить 0,5, так і несиметричною – імовірність випадання «герба» відмінна від 0,5; при цьому несиметричну й симетричну монети зовні розрізнити неможливо). Для цього підкидаємо монету n разів, n = 5, 10, 15,..., і реєструємо кількість «гербів», що випали.

Запропонувати правило (критерій) для дослідження монети на симетричність.

Яку мінімальну кількість разів необхідно підкинути монету, щоб симетрична монета бракувалася з імовірністю, не більшою ніж 0,1, а несиметрична, ймовірність випадання " герба" якої дорівнює 0,8, виявлялася з імовірністю 0,95?

Розв'язати поставлену задачу як задачу перевірки статистичних гіпотез (див. задачу 8).

Вказівка. За нульову вибрати гіпотезу: монета симетрична.

3. Діагностика початкової форми туберкульозу. Розглянемо задачу діагностики туберкульозу, пов'язану з установленням того факту, що метод рентгенівського аналізу не є абсолютно надійним при визначенні наявності або відсутності захворювання: здоровий індивідуум чи ні, результат рентгенівського дослідження на туберкульоз визначити наперед не можна.

З попереднього досвіду відомо, що:

а) якщо пацієнт хворий на туберкульоз, то ймовірність того, що окремий рентгенівський аналіз виявить хворобу (рентгенівський аналіз буде позитивним), становить p_0

б) якщо у пацієнта немає ніяких ознак туберкульозу, то ймовірність того, що окремий рентгенівський аналіз буде позитивним (у пацієнта буде виявлений туберкульоз), становить p_1.

Уявимо, що в певній клініці при обстеженні стану здоров'я пацієнта для виявлення можливих ознак туберкульозу одним і тим самим способом виготовляють n рентгенівських знімків. Тлумачення знімків здійснюється так, щоб була забезпечена незалежність діагнозу за кожним знімком від висновків, зроблених за іншими знімками.

Нехай в умовах описаного вище експерименту $p_0=0{,}75$, $p_1=0{,}05$, $n=3$ (виготовили три знімки). При цьому один з трьох аналізів виявився позитивним. Визначити, хворіє пацієнт на туберкульоз чи ні. Відповісти на це запитання, сформулювавши поставлену задачу як задачу перевірки статистичних гіпотез (див. задачу 8).

4. Про телепатів. Є люди, які стверджують, що вони можуть читати думки на відстані, тобто є телепатами (стверджують, що читають, але не гарантують, що правильно будуть прочитані всі думки, тобто визнають, що іноді помиляються: адже думки на відстані читати нелегко). Пропонується перевірити здібності телепата читати думки на відстані. З цією метою проведіть такий експеримент. Підкиньте правильну монету і зафіксуйте результат, наприклад на папері. Телепат читає вашу думку – «герб» чи «решка» – і також фіксує результат. Далі підраховується кількість правильно прочитаних «гербів» та «решок». Якщо телепат дійсно читає ваші думки, то частка правильно прочитаних «гербів» і «решок», тобто частка успіхів, буде великою (успіх – правильно прочитаний символ, невдача – неправильно прочитаний символ).

Монета підкидається n разів. При цьому телепат правильно прочитав ξ символів. Якого висновку ви дійдете щодо здібностей телепата читати думки?

Запропонувати критерій, за допомогою якого можна було б виявляти телепатичні здібності.

Сформулювати і розв'язати поставлену задачу в термінах перевірки статистичних гіпотез (див. задачу 8).

З'ясувати, як змінюється функція потужності критерію залежно від рівня значущості та кількості n підкидань монети.

Вказівка. За нульову вибирати гіпотезу: телепат думки не читає.

Перевірити телепатичні здібності свого товариша за описаним вище експериментом.

5. Про статистика й експериментатора. Два студенти (одного будемо називати статистиком, іншого – експериментатором) мають монети: симетричну і несиметричну; ймовірність випадання «герба» останньої становить p ($p\neq 1/2$). Експериментатор і статистик домовляються про таке: експериментатор виходить і до сусідньої кімнати, вибирає одну з монет (яку – статистик не знає), підкидає її n разів і результат (кількість випадінь «герба») повідомляє статистику. Останній стверджує, що він може визначити, яку з двох монет підкидав експериментатор. Однак той ставить під сумнів ці здібності статистика і як аргумент висуває такі заперечення: оскільки результат (кількість випадінь «герба») передбачити неможливо, причому як для симетричної монети, так і для несиметричної ця кількість може бути будь-якою: 0,1,2,...,n, визначити, яка з монет підкидалася, неможливо.

Хто має рацію – статистик чи експериментатор?

Станьте на позицію статистика: сформулюйте поставлену задачу як задачу перевірки статистичних гіпотез (див. задачу 8). Що можна сказати стосовно можливості розрізнити симетричну і несиметричну монети? Нехай $n = 25$. Яким буде висновок, якщо «герб» випав 21 раз?

Вказівка. За нульову вибрати гіпотезу: монета симетрична.

6. Контроль на токсичність. Процес виробництва деякого медичного препарату досить складний, так що неістотні на перший погляд відхилення від технології можуть спричинити появу високотоксичних побічних домішок. Токсичність останніх може виявитися такою сильною, що навіть незначна їх кількість, яку не можна встановити за допомогою звичайного хімічного аналізу, небезпечна для людини, яка буде вживати цей препарат. Тому до реалізації партії препарату її досліджують на токсичність біологічними методами. Малі дози препарату вводять певній кількості піддослідних тварин і результати реєструють. Якщо препарат токсичний, то всі або майже всі тварини гинуть – кількість згубних наслідків $\xi(\omega)$ велика ($\xi(\omega) \geq l$). У противному разі кількість згубних наслідків $\xi(\omega)$ мала ($\xi(\omega) < l$), або, що те саме, кількість тварин, які залишилися живі, велика.

Відомо, що для токсичного препарату ймовірність згубного наслідку не менша ніж 0,9, а якщо препарат нетоксичний, то ймовірність такого наслідку не перевищує 0,05. Яку мінімальну кількість n тварин слід ін’єктувати, щоб токсичний препарат виявлявся (класифікувався як токсичний) з імовірністю, не меншою ніж 0,999, а нетоксичний препарат витримував контроль (класифікувався як нетоксичний) з імовірністю 0,98?

Знайти значення n, сформулювавши поставлену задачу як задачу перевірки статистичних гіпотез (див. задачу 8).

7. Вибірковий контроль. Якість партії виробів вважається задовільною, якщо дефектні вироби в ній становлять не більше, ніж 2%. Яку кількість виробів партії необхідно випробувати для того, щоб партія, яка містить 8% дефектних виробів, відхилялася з імовірністю, не меншою ніж 0,95, а партія, що містить 2% таких виробів, приймалася з імовірністю, не меншою ніж 0,9?

Сформулювати та розв’язати поставлену задачу в термінах перевірки статистичних гіпотез (див. задачу 8).

Обчислити значення функції потужності критерію в точках 0,04; 0,06; 0,1 та ймовірність відхилення партії, що містить 4; 6; 10% дефектних виробів.

8. Про телепатів. Є люди, які стверджують, що вони можуть читати думки на відстані, тобто є телепатами. При цьому телепат не претендує на те. шо він безпомилково читає думки, але стверджує, що, іноді помиляючись, він усе-таки читає думки частіше правильно, ніж неправильно. Треба перевірити здібності телепата читати думки на відстані. З цією метою пропонується провести такий експеримент. Задумайте навмання число (для простоти 0 або 1) і зафіксуйте його, наприклад, на папері. Телепат читає вашу думку – задумане число і також фіксує його. Якщо телепат справді читає ваші думки, то частка правильно прочитаних нулів (нуль читається як нуль) і одиниць (одиниця читається як одиниця), тобто частка успіхів (успіх – правильно прочитаний символ) буде великою. Ви пропонуєте телепату прочитати n символів, $\xi(\omega)$ з них були прочитані правильно. Чи свідчить це про здібності телепата читати думки на відстані?

Запропонувати правило (критерій), за яким можна виявити здібності телепата читати думки на відстані.

Сформулювати і розв'язати поставлену задачу в термінах перевірки статистичних гіпотез, а саме:

сформулювати нульову (основну) й альтернативні гіпотези;

з'ясувати, в чому полягають помилки першого та другого родів;

визначити випадкову величину, яка спостерігається в експерименті (що означають висунуті гіпотези стосовно розподілів цієї випадкової величини?);

запропонувати критерій для перевірки нульової гіпотези;

призначити рівень значущості критерію, обґрунтувати свій вибір;

обчислити ймовірність помилки першого роду;

дослідити поведінку функції потужності критерію;

дати частотну інтерпретацію одержаних результантів.

9. Контроль чистоти води. Для потреб деякого хімічного виробництва необхідно, щоб вода, яка використовується, була чистою, тобто містила мало бактерій. Вода, на одиницю об'єму якої в середньому припадає менше двох бактерій, вважається придатною для зазначеного хімічного виробництва. Якщо ж середня кількість бактерій на одиницю об'єму води дорівнює двом або більше, то міра її забруднення недопустима. Звичайна методика контролю води на чистоту така. Береться 10 (у загальному випадку n) проб води одиничного об'єму. Потім кожну з цих проб додають у колбу з живильним середовищем, яку після цього тримають при температурі, сприятливій для росту бактерій. Якщо проба забруднена, тобто містить принаймні одну бактерію, то колонія бактерій росте і розчин, який спочатку був прозорий, стає каламутним. Про чистоту води судять за кількістю ξ забруднених проб: якщо ця кількість не перевищує l, то вода вважається чистою; у противному разі провадиться додаткове очищення води, що, природно, потребує праці та витрат, і вона знову досліджується на чистоту. Однак, якщо вода, що використовується, містить значну кількість бактерій, то спричинені цим збитки значно більші.

Яким має бути l, щоб вода з середнім умістом бактерій – дві на одиницю об'єму – виявлялася з імовірністю 0,99? Знайти значення l, формулюючи поставлену задачу як задачу перевірки статистичних гіпотез (див. задачу 8).

Зауваження. За розподіл кількості бактерій в одиничному об'ємі води з середнім умістом λ бактерій на одиницю об'єму природно прийняти пуассонівський розподіл із параметром λ.

10. Про пані, яка куштує чай. Одна пані стверджує, що, покуштувавши чашку чаю з молоком, вона може визначити, що було спочатку налите в чашку – молоко чи чай. При цьому пані не претендує на безпомилкове визначення різниці на смак, але стверджує, що, нехай іноді помиляючись, вона частіше визначає правильно, ніж неправильно.

Перш ніж визнати здібності пані, їй пропонують взяти участь в такому експерименті. Пані необхідно покуштувати і класифікувати n пар чашок чаю (по парі дат шок за сніданком протягом n днів). У кожну пару входить по чашці чаю, що готувався за різними рецептами. Кількість ξ правильно класифікованих пар реєструється. Нехай пані з n пар чашок правильно класифікувала $\xi(\omega)$ пар. Чи свідчить це про наявність у неї здібностей розрізняти рецепти, за якими готується чай?

Ви – член доброзичливого журі – не хочете несправедливо нехтувати здібностями пані, коли вона справді їх має. Запропонуйте правило (критерій), яким варто керуватися журі, роблячи висновок про здібності пані розрізняти рецепти приготування чаю.

Сформулюйте та розв'яжіть поставлену задачу в термінах перевірки статистичних гіпотез (див. задачу 8).

З'ясуйте, як змінюється функція потужності критерію залежно від рівня його значущості та кількості n пар чашок чаю, що пропонується для класифікації.

Зауваження. Як член журі ви, зрозуміло, не хочете визнати вищезгадані здібності пані, якщо вона їх не має; тому за нульову гіпотезу пропонуєте гіпотезу H_0: пані не має здібностей. Але як член доброзичливого журі ви не хочете несправедливо нехтувати здібностями пані, якщо вона їх має; тому ви не призначаєте занадто малий рівень значущості критерію.

11. Вибірковий контроль. Споживач купує великі партії товарів. Щоб уникнути значної частки дефектних виробів, він вибирає для контролю з кожної партії n виробів. Партія приймається, якщо частка дефектних виробів не перевищує l.

Споживач висуває такі вимоги до контролю: якщо частка дефектних виробів у партії становить 8%, то партія має бракуватися з імовірністю, не меншою ніж 0,9. У постачальника свої вимоги: він хоче, щоб партія, частка дефектних виробів у якій становить 1%, бракувалася з імовірністю, не більшою ніж 0,04.

Якими мають бути обсяг (зрозуміло, мінімальний) вибірки n з партії та значення l?

Сформулювати і розв'язати поставлену задачу в термінах перевірки статистичних гіпотез (див. задачу 8).

Яка ймовірність забракувати партію товарів, що містить 1; 2; 4; 8; 10% дефектних виробів?

12. Про гральні кубики. Маємо пару гральних кубиків: один – симетричний (імовірність випадання кожної грані дорівнює 1/6), інший – несиметричний, центр тяжіння якого зміщений так, що частіше з'являються 6; 5; 4 очок. Кубики зовні нерозрізненні.

Щоб виявити несиметричний кубик, беремо один із кубиків (який саме невідомо – їх зовні розрізнити неможливо) і підкидаємо його n разів, реєструючи кількість очок, що випали. Яку мінімальну кількість разів необхідно підкинути кубик, щоб симетричний бракувався з імовірністю, не більшою ніж 0,1, а несиметричний, з імовірністю 0,9 випадання кількості очок більшої 3, виявлявся з імовірністю 0,95?

Розв'язати поставлену задачу, сформулювавши її як задачу перевірки статистичних гіпотез (див. задачу 8).

З'ясувати, як змінюється функція потужності критерію залежно від рівня його значущості та числа підкидань кубика в експерименті.

Обчислити ймовірності того, що несиметричний кубик, імовірність випадання кількості очок більшої ніж 3 якого становить 0,6; 0,7; 0,8; 0,9, буде виявлений; не буде виявлений.

Вказівка. За нульову вибрати гіпотезу: гральний кубик симетричний.

13. Розв'язати задачу 9 за умови, що в експерименті використовується n проб води.

З'ясувати, як змінюється функція потужності критерію залежно від рівня його значущості та кількості проб води, що використовується в експерименті.

14. В умовах експерименту, описаного в задачі 12, кубик підкидають 15 разів. За результатами підкидань – кількістю очок, що випали, – зробити висновок про симетричність грального кубика. При цьому симетричний кубик має бракуватися з імовірністю, не більшою ніж 0,05.

Як зміниться висновок, якщо кубик підкинути 30 разів?

Сформулювати та розв'язати поставлену задачу в термінах перевірки статистичних гіпотез (див. задачу 8).

Обчислити ймовірності того, що несиметричний кубик, імовірність випадання кількості очок більшої ніж 3 якого становить 0,6; 0,7; 0,8; 0,9, буде виявлений; не буде виявлений.

15. Булочки з ізюмом. Державним стандартом установлено, що при випіканні солодких булочок на 1000 виробів має припадати 10 000 штук ізюму (в середньому 10 штук на одну булочку). У нас, однак, є сумнів, що весь ізюм використано за призначенням (він міг, принаймні частково, розійтися іншими каналами), і ми хочемо перевірити, чи це так. Для цього купуємо одну булочку й рахуємо кількість ізюму в ній: виявилося, що в ній є $\xi(\omega)$ штук ізюму. За кількістю $\xi(\omega)$ ізюму в булочці треба зробити висновок, чи весь він використовується за призначенням.

Розв'язати поставлену задачу, сформулювавши її як задачу перевірки статистичних гіпотез (див. задачу 8).

Як зміниться висновок, коли, купивши дві булочки, порахувати кількість ізюму в них?

16. Контроль на токсичність. Медичний препарат проходить контроль на токсичність за методикою, описаною в задачі 6.

Відомо, що ймовірність летального кінця від використання токсичного препарату не менша ніж 0,8, а якщо препарат нетоксичний, то ймовірність летального кінця не перевищує 0,05.

З'ясувати, чи є досліджуваний препарат токсичним, якщо його ввели 10 мишам і при цьому було зафіксовано загибель двох тварин.

Розв'язати поставлену задачу, сформулювавши її як задачу перевірки статистичних гіпотез (див. задачу 8).

17. Таблиці випадкових чисел. Таблиці, що пропонуються як таблиці випадкових чисел, мають задовольняти низку вимог випадковості. Зокрема, ймовірність появи, наприклад, числа 99 чи будь-якого іншого двозначного числа в них має становити 1/100. Щоб упевнитися, що таблиця випадкових чисел цю вимогу задовольняє, виберемо з неї 400 двозначних чисел, починаючи з будь-якого місця таблиці, і підрахуємо кількість появ ξ числа 99 серед них. Чи узгоджується твердження про розглянуту таблицю як про таблицю випадкових чисел з результатами експерименту?

Сформулювати і розв'язати поставлену задачу в термінах перевірки статистичних гіпотез (див. задачу 8).

Вказівка. Якщо таблиця справді є таблицею випадкових чисел, то випадкова величина ξ – кількість появ числа 99 серед 400 двозначних чисел – має біномний розподіл з параметрами 400 та 1/100. Тому у зв'язку з описаним експериментом за основну гіпотезу природно взяти гіпотезу H_0: ξмає біномний розподіл з параметрами 400 і 1/100, а саме:

$$\mathrm{P}\{\xi = k\} = \mathrm{C}_{400}^{k}\left(\frac{1}{100}\right)^{k}\left(1-\frac{1}{100}\right)^{400-k},\ k = 0,1,\ldots,400.$$

Отже, перевірка таблиці на випадковість за описаним експериментом зводиться до перевірки гіпотези H_0.

Зазначимо, що, розв'язуючи поставлену задачу, біномний розподіл зручно апроксимувати пуассонівським.

Зауваження. Таблицю випадкових чисел сформуйте на листі електронної таблиці за допомогою формули =ROUND(RAND()*400;0).

18. Маємо п'ять монет, імовірність випадіння «герба» яких становить відповідно 0,5; 0,6; 0,7; 0,8; 0,9. Монети за зовнішнім виглядом розрізнити неможливо.

Щоб виявити симетричну монету, беремо одну з них (яку саме невідомо, оскільки зовні вони нерозрізненні) й підкидаємо її 20 разів.

За результатами підкидань зробити висновок про симетричність монети.

Сформулювати і розв'язати поставлену задачу в термінах перевірки статистичних гіпотез (див. задачу 8), вибравши як основну гіпотезу: монета симетрична. Побудувати критерій для перевірки цієї гіпотези з рівнем значущості $\alpha = 0{,}05$, а також функцію потужності критерію. Яка ймовірність виявлення несиметричної монети, якщо ймовірність випадання її «герба» становить 0,6; 0,7; 0,8; 0,9?

19. Якість інсектициду. Як фахівця з математичною підготовкою вас запрошують проаналізувати результати дослідницької роботи, пов'язаної з виробництвом інсектицидів. Якість інсектициду визначається відсотком ураження – відсотком комах, які гинуть у групі, обробленій інсектицидом. Кращий з наявних інсектицидів має відсоток ураження 92. В лабораторії створено новий інсектицид ШС. Виникає запитання: чи є він якіснішим?

Щоб з'ясувати це, пропонується обробити новим інсектицидом 100 комах і підрахувати кількість тих, які вижили. Якщо кількість таких комах менша від деякого числа l, то ШС якісніший, в противному разі – ні.

Автори нового інсектициду стверджують, що відсоток ураження нового препарату не менший 98, і хочуть, щоб ця його властивість проявлялася з імовірністю, не нижчою ніж 0,94. Інсектицид ШС дорожчий за наявні в продажу препарати, і споживач, природно, відмовляється купувати ШС, якщо його відсоток ураження 92, наполягаючи при цьому, щоб останнє проявлялося з імовірністю 0,95.

Чи можна за кількістю комах, які вижили (зі 100), зробити висновок про якість інсектициду ШС, забезпечивши перелічені вище умови? Якщо 100 комах у досліді замало, то якою має бути мінімальна кількість комах в оброблюваній групі, щоб можна було відповісти на поставлене запитання?

Відповісти на ці запитання, сформулювавши поставлену задачу як задачу перевірки статистичних гіпотез (див. задачу 8).

20. Випадково – не довільно. Випадково не означає довільно. Випадковість підпорядковується своїм суворим законам (твердження на перший погляд парадоксальне). Запропонувати випадкову послідовність чисел може далеко не кожний.

Перевірте свої здібності. Випишіть послідовність із 500 випадкових чисел, наприклад двозначних. Якщо запропонована вами послідовність чисел випадкова, то ймовірність появи, наприклад, числа 77 (чи будь-якого іншого двозначного) становить 0,01. Перевірте, чи задовольняє виписана вами послідовність цю вимогу.

Сформулюйте й розв'яжіть поставлену задачу як задачу перевірки статистичних гіпотез (див. задачу 8).

Вказівка. Якщо в запропонованій послідовності чисел імовірність того, що з'явиться число 77, становить 0,01, то випадкова величина ξ – кількість появ числа 77 в послідовності з 500 двозначних чисел – має біномний розподіл

$$\mathrm{P}\{\xi = k\} = C_{500}^{k}\left(\frac{1}{100}\right)^{k}\left(1 - \frac{1}{100}\right)^{500-k},$$

де $k = 0,1,...,500$, а враховуючи, що $n = 500$ – кількість випробувань – велика, а p=0,01 – імовірність успіху – мала, можна вважати, що ξ має пуассонівський розподіл з параметром $\lambda=np=500\cdot0{,}01=5$. Тому для перевірки того, що ймовірність появи числа 77 становить 0,01, стосовно розподілу випадкової величини ξ висуваємо гіпотезу H: ξ має пуассонівський розподіл із параметром $\lambda = 5$ і перевіряємо її, реєструючи значення ξ (підраховуючи кількість появ числа 77 серед 500 запропонованих).

21. Якість інсектициду. Ви займаєтесь дослідницькою роботою, пов'язаною з виробництвом інсектицидів. Якість інсектициду визначається відсотком ураження – відсотком комах, які гинуть від препарату в групі, що досліджується. Кращий з наявних інсектицидів має відсоток ураження 90. В лабораторії створено новий інсектицид NIP. Виникає запитання: чи є цей препарат кращим?

Щоб з'ясувати, чи є різниця в якості інсектицидів, група з n комах обробляється NIP і підраховується кількість комах, які вижили. Якщо їх менше певного числа l, то вважається, що інсектицид NIP якісніший, у противному разі – ні.

Автори нового препарату стверджують, що він має відсоток ураження, не менший ніж 99, і хочуть, щоб це виявлялося з імовірністю, не меншою ніж 0,98. Інсектицид NIP дорожчий за наявні в продажу препарати, і споживач, природно, відмовляється купувати NIP, якщо його відсоток ураження 90, наполягаючи, щоб останнє проявлялося з імовірністю 0,95.

Треба вибрати n (якомога меншим) та l так, щоб забезпечити перелічені вище вимоги.

Сформулювати і розв'язати поставлену задачу як задачу перевірки статистичних гіпотез (див. задачу 8).

22. Діагностика початкової форми туберкульозу. Нехай в умовах експерименту, описаного в задачі 3, значення p_0= 0,9, p_1= 0,01.

Яку мінімальну кількість n знімків необхідно зробити при обстеженні кожного пацієнта, щоб пацієнт з ознаками туберкульозу виявлявся з імовірністю 0,99, а пацієнт, який не має таких ознак, витримував перевірку на туберкульоз з імовірністю 0,95.

Визначити n (та відповідне значення l), сформулювавши поставлену задачу як задачу перевірки статистичних гіпотез (див. задачу 8).