Добрый день, рассмотрите и законспектируйте данный материал.

Ответьте на вопросы:

1. Что такое C_n^m – **биномиальный коэффициент**;
2. Что такое **факториал.**
3. Что такое **статистическая вероятность.**
4. Что такое Если вероятность p появления случайного события A в каждом испытании постоянна, то вероятность $P_n(m)$ того, что в n испытаниях событие A наступит ровно m раз, приближённо равна:
$$P_n(m) \approx \frac{1}{\sqrt{npq}} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} = \frac{1}{\sqrt{npq}} \cdot \varphi(x)$$
, где $x = \frac{m - np}{\sqrt{npq}}$.
5. Что такое Если вероятность p появления случайного события A в каждом испытании постоянна, то вероятность $P_n(m_1 \le m \le m_2)$ того, что в n испытаниях событие A наступит **не менее** m_1 **и не более** m_2 **раз** *(от* m_1 *до* m_2 *раз включительно)*, приближённо равна:

$$P_n(m_1 \le m \le m_2) \approx \Phi(x_2) - \Phi(x_1)$$
, где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_0^x e^{-\frac{z^2}{2}} dz, \quad x_2 = \frac{m_2 - np}{\sqrt{npq}}, \quad x_1 = \frac{m_1 - np}{\sqrt{npq}}$

6. Что такое **функцией Лапласа.**
7. Что такое *интервал.*

Локальная и интегральная теоремы Лапласа

Данная статья является естественным продолжением урока о **независимых испытаниях**, на котором мы познакомились с **формулой Бернулли** и отработали типовые примеры по теме. Локальная и интегральная теоремы Лапласа (Муавра-Лапласа) решают аналогичную задачу с тем отличием, что они применимы к достаточно большому количеству независимых испытаний. Не нужно тушеваться слов «локальная», «интегральная», «теоремы» – материал осваивается с той же лёгкостью, с какой Лаплас потрепал кучерявую голову Наполеона. Поэтому безо всяких комплексов и предварительных замечаний сразу же рассмотрим демонстрационный пример:

Монета подбрасывается 400 раз. Найти вероятность того, что орёл выпадет 200 раз.

По характерным признакам здесь следует применить **формулу Бернулли** $P_n^m = C_n^m p^m q^{n-m}$. Вспомним смысл этих букв:

P_n^m – вероятность того, что в n независимых испытаниях случайное событие A наступит ровно m раз;

C_n^m – **биномиальный коэффициент**;
p – вероятность появления события A в каждом испытании;
$q = 1 - p$ – вероятность противоположного события.

Применительно к нашей задаче:
$n = 400$ – общее количество испытаний;
$m = 200$ – количество бросков, в которых должен выпасть орёл;
$p = 0{,}5$ – вероятность выпадения орла в каждом броске;
$q = 1 - p = 0{,}5$ – вероятность выпадения решки.

Таким образом, вероятность того, что в результате 400 бросков монеты орёл выпадет ровно 200 раз: $P_{400}^{200} = C_{400}^{200} \cdot (0{,}5)^{200} \cdot (0{,}5)^{200} = \frac{400!}{200! \cdot 200!} \cdot (0{,}5)^{400}$ …Стоп, что делать дальше? Микрокалькулятор (по крайне мере, мой) не справился с 400-й степенью и капитулировал перед **факториалами**. А считать через произведение что-то не захотелось =) Воспользуемся **стандартной функцией Экселя**, которая сумела обработать монстра: $P_{400}^{200} = 0{,}0398693019637926$.

Заостряю ваше внимание, что получено **точное** значение и такое решение вроде бы идеально. На первый взгляд. Перечислим веские контраргументы:

– во-первых, программного обеспечения может не оказаться под рукой;
– и во-вторых, решение будет смотреться нестандартно *(с немалой вероятностью придётся перерешивать)*;

Поэтому, уважаемые читатели, в ближайшем будущем нас ждёт:

Локальная теорема Лапласа

Если вероятность p появления случайного события A в каждом испытании постоянна, то вероятность $P_n(m)$ того, что в n испытаниях событие A наступит ровно m раз, приближённо равна:

$$P_n(m) \approx \frac{1}{\sqrt{npq}} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} = \frac{1}{\sqrt{npq}} \cdot \varphi(x)$$, где $$x = \frac{m - np}{\sqrt{npq}}$$.

При этом, чем больше n, тем рассчитанная вероятность $P_n(m)$ будет лучше приближать точное значению P_n^m, полученное *(хотя бы гипотетически)* по формуле Бернулли. Рекомендуемое минимальное количество испытаний – примерно 50-100, в противном случае результат $P_n(m)$ может оказаться далёким от истины. Кроме того, локальная теорема Лапласа работает тем лучше, чем вероятность p ближе к 0,5, и наоборот – даёт существенную погрешность при значениях p, близких к нулю либо единице. По этой причине ещё одним критерием эффективного использования формулы $P_n(m) \approx \frac{1}{\sqrt{npq}} \cdot \varphi(x)$ является выполнение неравенства $npq > 10$ (≈ 10).

Так, например, если $n = 50,\ p = 0{,}5$, то $npq = 50 \cdot 0{,}5 \cdot 0{,}5 = 12{,}5 > 10$ и применение теоремы Лапласа для 50 испытаний оправдано. Но если $n = 50$ и $p = 0{,}1$, то $npq = 50 \cdot 0{,}1 \cdot 0{,}9 = 4{,}5 < 10$ и приближение $P_n(m)$ *(к точному значению* P_n^m *)* будет плохим.

О том, почему $P_n(m) \approx P_n^m$ и об особенной функции $\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$ мы поговорим на уроке о **нормальном распределении вероятностей**, а пока нам потребуется формально-вычислительная сторона вопроса. В частности, важным фактом является **чётность** этой функции: $\varphi(-x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(-x)^2}{2}} = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} = \varphi(x)$.

Оформим официальные отношения с нашим примером:

<u>Задача 1</u>

Монета подбрасывается 400 раз. Найти вероятность того, что орёл выпадет ровно:

а) 200 раз;
б) 225 раз.

С чего начать **решение**? Сначала распишем известные величины, чтобы они были перед глазами:

$n = 400$ – общее количество независимых испытаний;
$p = 0{,}5$ – вероятность выпадения орла в каждом броске;
$q = 1 - p = 0{,}5$ – вероятность выпадения решки.

а) Найдём вероятность того, что в серии из 400 бросков орёл выпадет ровно $m = 200$ раз. Ввиду большого количества испытаний используем локальную теорему Лапласа: $P_n(m) \approx \frac{1}{\sqrt{npq}} \cdot \varphi(x)$,

где $\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}, \quad x = \frac{m - np}{\sqrt{npq}}$.

На первом шаге вычислим требуемое значение аргумента:

$$x = \frac{m - np}{\sqrt{npq}} = \frac{200 - 400 \cdot 0{,}5}{\sqrt{400 \cdot 0{,}5 \cdot 0{,}5}} = \frac{200 - 200}{\sqrt{100}} = \frac{0}{10} = 0$$

Далее находим соответствующее значение функции: $\varphi(0)$. Это можно сделать несколькими способами. В первую очередь, конечно же, напрашиваются непосредственные вычисления:

$$\varphi(0) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{0^2}{2}} = \frac{1}{\sqrt{2\pi}} \cdot e^0 = \frac{1}{\sqrt{2\pi}} \approx 0{,}3989$$

Округление проводят, как правило, до 4 знаков после запятой.

Недостаток прямого вычисления состоит в том, что экспоненту переваривает далеко не каждый микрокалькулятор, кроме того, расчёты не особо приятны и отнимают время. Зачем так мучиться? Используйте **калькулятор по терверу** *(пункт 4)* и получайте значения $\varphi(x)$ моментально!

Кроме того, существует *таблица значений функции* $\varphi(x)$, которая есть практически в любой книге по теории вероятностей, в частности, в учебном пособии *В.Е. Гмурмана*. Закачайте, кто ещё не закачал – там вообще много полезного ;-) **И обязательно научитесь пользовать таблицей (прямо сейчас!)** – подходящей вычислительной техники всегда может не оказаться под рукой!

На заключительном этапе применим формулу $P_n(m) \approx \frac{1}{\sqrt{npq}} \cdot \varphi(x)$:

$P_{400}(200) \approx \frac{1}{10} \cdot \varphi(0) \approx 0{,}1 \cdot 0{,}3989 = 0{,}03989$ – вероятность того, что при 400 бросках монеты орёл выпадет ровно 200 раз.

Как видите, полученный результат очень близок к точному значению $P_{400}^{200} = 0{,}0398693019637926$, вычисленному по **формуле Бернулли**.

б) Найдём вероятность того, что в серии из 400 испытаний орёл выпадет ровно $m = 225$ раз. Используем локальную теорему Лапласа. Раз, два, три – и готово:

$$x = \frac{m - np}{\sqrt{npq}} = \frac{225 - 200}{10} = \frac{25}{10} = 2{,}5$$

$$\varphi(2{,}5) \approx 0{,}0175$$

$P_{400}(225) \approx \frac{1}{10} \cdot \varphi(2{,}5) \approx 0{,}1 \cdot 0{,}0175 = 0{,}00175$ – искомая вероятность.

Ответ: а) $\approx 0{,}04$; б) $\approx 0{,}002$

Следующий пример, как многие догадались, посвящён деторождению – и это вам для самостоятельного решения :)

Задача 2

Вероятность рождения мальчика равна 0,52. Найти вероятность того, что среди 100 новорожденных окажется ровно: а) 40 мальчиков, б) 50 мальчиков, в) 30 девочек.

Результаты округлить до 4 знаков после запятой.

…Интересно тут звучит словосочетание «независимые испытания» =) Кстати, реальная **статистическая вероятность** рождения мальчика во многих регионах мира колеблется в пределах от 0,51 до 0,52.

Примерный образец оформления задачи в конце урока.

Все заметили, что числа получаются достаточно малыми, и это не должно вводить в заблуждение – ведь речь идёт о вероятностях отдельно взятых, *локальных* значениях (отсюда и название теоремы). А таковых значений много, и, образно говоря, вероятности «должно хватить на всех». Правда, многие события будут **практически невозможными**.

Поясню вышесказанное на примере с монетами: в серии из четырёхсот испытаний орёл теоретически может выпасть от 0 до 400 раз, и данные события образуют **полную группу**:
$P_{400}(0)+P_{400}(1)+P_{400}(2)+\ldots+P_{400}(199)+P_{400}(200)+P_{400}(201)+\ldots+P_{400}(399)+P_{400}(400)=1$

Однако бОльшая часть этих значений представляет собой сущий мизер, так, например, вероятность того, что орёл выпадет 250 раз – уже одна десятимиллионная: $P_{400}(250)\approx 0{,}0000001$. О значениях наподобие $P_{400}(100),\ P_{400}(350)$ тактично умолчим =)

С другой стороны, не следует недооценивать и скромные результаты: если $P_{400}(225)$ составляет всего около $0{,}00175$, то вероятность того, орёл выпадет, скажем, *от 220 до 250 раз*, будет весьма заметна.

А теперь задумаемся: как вычислить данную вероятность? Не считать же по **теореме сложения вероятностей несовместных событий** сумму:
$P_{400}(220)+P_{400}(221)+P_{400}(222)+\ldots+P_{400}(249)+P_{400}(250)$

Гораздо проще эти значения *объединить*. А объединение чего-либо, как вы знаете, называется **интегрированием**:

Интегральная теорема Лапласа

Если вероятность p появления случайного события A в каждом испытании постоянна, то вероятность $P_n(m_1\le m\le m_2)$ того, что в n испытаниях событие A наступит **не менее** m_1 **и не более** m_2 **раз** *(от* m_1 *до* m_2 *раз включительно)*, приближённо равна:

$P_n(m_1\le m\le m_2)\approx\Phi(x_2)-\Phi(x_1)$, где $\Phi(x)=\frac{1}{\sqrt{2\pi}}\cdot\int_0^x e^{-\frac{z^2}{2}}dz,\quad x_2=\frac{m_2-np}{\sqrt{npq}},\quad x_1=\frac{m_1-np}{\sqrt{npq}}$

При этом количество испытаний, разумеется, тоже должно быть достаточно большим и вероятность p не слишком мала/велика *(ориентировочно* $npq>10$*)*, иначе приближение будет неважным либо плохим.

Функция $\Phi(x)$ называется **функцией Лапласа**, и её значения опять же сведены в стандартную таблицу (**найдите и научитесь с ней работать!!**). Микрокалькулятор здесь не поможет, поскольку интеграл является неберущимся. Но вот в Экселе есть соответствующий функционал – используйте *пункт 5* **расчётного макета**.

На практике наиболее часто встречаются следующие значения:
$\Phi(0)=0;\quad \Phi(1)\approx 0{,}3413;\quad \Phi(2)\approx 0{,}4772;\quad \Phi(3)\approx 0{,}4987$ – перепишите к себе в тетрадь.

Начиная с $x = 4$, можно считать, что $\Phi(x) = 0{,}5$, или, если записать строже: $\lim\limits_{x\to+\infty} \Phi(x) = 0{,}5$

Кроме того, функция Лапласа **нечётна**: $\Phi(-x) = -\Phi(x)$, и данное свойство активно эксплуатируется в задачах, которые нас уже заждались:

Задача 3

Вероятность поражения стрелком мишени равна 0,7. Найти вероятность того, что при 100 выстрелах мишень будет поражена от 65 до 80 раз.

Я подобрал наиболее реалистичный пример, а то у меня тут нашлось несколько задач, в которых стрелок делает тысячи выстрелов =)

Решение: в данной задаче речь идёт о **повторных независимых испытаниях**, причём их количество достаточно велико. По условию требуется найти вероятность того, что мишень будет поражена не менее 65, но и не более 80 раз, а значит, нужно использовать интегральную теорему Лапласа: $P_n(m_1 \le m \le m_2) \approx \Phi(x_2) - \Phi(x_1)$,

где $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \cdot \displaystyle\int_0^x e^{-\frac{z^2}{2}} dz, \quad x_2 = \dfrac{m_2 - np}{\sqrt{npq}}, \quad x_1 = \dfrac{m_1 - np}{\sqrt{npq}}$

Для удобства перепишем исходные данные в столбик:
$n = 100$ – всего выстрелов;
$m_1 = 65$ – минимальное число попаданий;
$m_2 = 80$ – максимальное число попаданий;
$p = 0{,}7$ – вероятность попадания в мишень при каждом выстреле;
$q = 1 - p = 0{,}3$ – вероятность промаха при каждом выстреле.

$npq = 100 \cdot 0{,}7 \cdot 0{,}3 = 21 > 10$, следовательно, теорема Лапласа даст хорошее приближение.

Вычислим значения аргументов:

$$x_2 = \frac{m_2 - np}{\sqrt{npq}} = \frac{80 - 100 \cdot 0{,}7}{\sqrt{21}} = \frac{80 - 70}{\sqrt{21}} \approx \frac{10}{4{,}5825} \approx 2{,}18$$

$$x_1 = \frac{m_1 - np}{\sqrt{npq}} = \frac{65 - 70}{\sqrt{21}} \approx \frac{-5}{4{,}5825} \approx -1{,}09$$

Обращаю ваше внимание, что произведение npq вовсе не обязано нацело извлекаться из-под корня *(как любят «подгонять» числа авторы задач)* – без тени сомнения извлекаем корень и округляем результат; я привык оставлять 4 знака после запятой. А вот полученные значения x_1, x_2 обычно округляют до 2 знаков после запятой – эта традиция идёт из *таблицы значений функции* $\Phi(x)$, где аргументы представлены именно в таком виде.

Используем указанную выше таблицу либо **расчётный макет по терверу** *(пункт 5)*.
В качестве письменного комментария советую поставить следующую фразу: *значения функции* $\Phi(x)$ *найдём по соответствующей таблице*:

$P_{100}(65 \le m \le 80) \approx \Phi(2{,}18) - \Phi(-1{,}09) = \Phi(2{,}18) + \Phi(1{,}09) = 0{,}4854 + 0{,}3621 = 0{,}8475$
– вероятность того, что при 100 выстрелах мишень будет поражена от 65 до 80 раз.

Обязательно пользуемся нечётностью функции! На всякий случай распишу подробно:
$\Phi(2{,}18) - \Phi(-1{,}09) = \Phi(2{,}18) - (-\Phi(1{,}09)) = \Phi(2{,}18) + \Phi(1{,}09)$

Дело в том, что *таблица значений функции* $\Phi(x)$ содержит только положительные «икс», а мы работаем *(по крайне мере, по «легенде»)* с таблицей!

Ответ: $P_{100}(65 \le m \le 80) \approx 0{,}8475$

Результат чаще всего округляют до 4 знаков после запятой *(опять же в соответствии с форматом таблицы)*.

Для самостоятельного решения:

Задача 4

В здании имеется 2500 ламп, вероятность включения каждой из них в вечернее время равна 0,5. Найти вероятность того, что вечером будет включено не менее 1250 и не более 1275 ламп.

Примерный образец чистового оформления в конце урока.

Следует отметить, что рассматриваемые задачи очень часто встречаются в «обезличенном» виде, например:

Производится некоторый опыт, в котором случайное событие A может появиться с вероятностью 0,5. Опыт повторяется в неизменных условиях 2500 раз. Определить вероятность того, что в 2500 опытах событие A произойдет от 1250 до 1275 раз

И подобных формулировок выше крыши. По причине трафаретности задач условие нередко стремятся завуалировать – это «единственный шанс» хоть как-то разнообразить и усложнить решение:

Задача 5

В институте обучается 1000 студентов. В столовой имеется 105 посадочных мест. Каждый студент отправляется в столовую на большой перемене с вероятностью 0,1. Какова вероятность того, что в обычный учебный день:

а) столовая будет заполнена не более чем на две трети;
б) посадочных мест на всех не хватит.

Обращаю внимание на существенную оговорку «в ОБЫЧНЫЙ учебный день» – она обеспечивает относительную неизменность ситуации. После праздников в институт может прийти значительно меньше студентов, а на «День открытых дверей» нагрянуть голодная делегация =) То есть, в «необычный» день вероятности будут заметно отличаться.

Решение: используем интегральную теорему Лапласа $P_n(m_1 \le m \le m_2) \approx \Phi(x_2) - \Phi(x_1)$,

где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_0^x e^{-\frac{z^2}{2}} dz, \quad x_2 = \frac{m_2 - np}{\sqrt{npq}}, \quad x_1 = \frac{m_1 - np}{\sqrt{npq}}$

В данной задаче:

$n = 1000$ – всего студентов в институте;

$p = 0{,}1$ – вероятность того, что студент отправится в столовую на большой перемене;

$q = 1 - p = 1 - 0{,}1 = 0{,}9$ – вероятность противоположного события.

а) Вычислим, сколько посадочных мест составляют две трети от общего количества: $105 \cdot \frac{2}{3} = 70$ мест

Найдём вероятность того, что в обычный учебный день столовая будет заполнена не более чем на две трети. Что это значит? Это значит, что на большой перемене придут от 0 до 70 человек. То, что никто не придёт или придут всего несколько студентов – есть события **практически невозможные**, однако в целях применения интегральной теоремы Лапласа эти вероятности все равно следует учесть. Таким образом:

$m_1 = 0, m_2 = 70$

Вычислим соответствующие аргументы:

$$x_2 = \frac{m_2 - np}{\sqrt{npq}} = \frac{70 - 1000 \cdot 0{,}1}{\sqrt{1000 \cdot 0{,}1 \cdot 0{,}9}} = \frac{70 - 100}{\sqrt{90}} \approx \frac{-30}{9{,}4868} \approx -3{,}16$$

$$x_1 = \frac{m_1 - np}{\sqrt{npq}} = \frac{0 - 100}{\sqrt{90}} \approx \frac{-100}{9{,}4868} \approx -10{,}54$$

В результате:

$$P_{1000}(0 \le m \le 70) \approx \Phi(-3{,}16) - \Phi(-10{,}54) = -\Phi(3{,}16) + \Phi(10{,}54) \approx$$

$\approx -0{,}4992 + 0{,}5 = 0{,}0008$ – вероятность того, что в обычный учебный день столовая будет заполнена не более чем на две трети.

Напоминание*: при* $x \ge 4$ *функцию Лапласа считаем равной* $\Phi(x) = 0{,}5$.

Толкучка, однако =)

б) Событие *«Посадочных мест на всех не хватит»* состоит в том, что в столовую на большой перемене придут обедать от 106 до 1000 человек *(главное, хорошо уплотнить =))*. Понятно, что высокая посещаемость невероятна, но тем не менее: $m_1 = 106, m_2 = 1000$.

Рассчитываем аргументы:

$$x_2 = \frac{m_2 - np}{\sqrt{npq}} = \frac{1000 - 100}{\sqrt{90}} \approx \frac{900}{9{,}4868} \approx 94{,}87$$

$$x_1 = \frac{m_1 - np}{\sqrt{npq}} = \frac{106 - 100}{\sqrt{90}} \approx \frac{6}{9{,}4868} \approx 0{,}63$$

Таким образом, вероятность того, что посадочных мест на всех не хватит:
$P_{1000}(106 \le m \le 1000) \approx \Phi(94,87) - \Phi(0,63) \approx 0,5 - 0,2357 = 0,2643$

Ответ: а) $\approx 0,0008$; б) $\approx 0,2643$

А теперь остановимся на одном **важном нюансе** метода: когда мы проводим вычисления на отдельно взятом отрезке, то всё «безоблачно» – решайте по рассмотренному шаблону. Однако в случае рассмотрения **полной группы событий** следует проявить *определённую аккуратность*. Поясню этот момент на примере только что разобранной задачи. В пункте «бэ» мы нашли вероятность $P_{1000}(106 \le m \le 1000) \approx 0,2643$ – того, что посадочных мест на всех не хватит. Далее, по той же самой схеме рассчитаем:
$P_{1000}(0 \le m \le 105) \approx 0,7019$ – вероятность того, что мест хватит.

Поскольку эти события **противоположны**, то сумма вероятностей должна равняться единице:
$P_{1000}(0 \le m \le 105) + P_{1000}(106 \le m \le 1000) = 0,7019 + 0,2643 = 0,9662$

?! В чём дело? – вроде бы тут всё логично. Дело в том, что функция Лапласа является **непрерывной**, а мы не учли *интервал* от 105 до 106. Вот здесь то и пропал кусочек 0,0338. Поэтому по той же самой стандартной формуле следует вычислить:
$P_{1000}(0 \le m < 106) \approx 0,7357$

Ну, или ещё проще:
$P_{1000}(0 \le m < 106) = 1 - P_{1000}(106 \le m \le 1000) = 1 - 0,2643 = 0,7357$

Возникает вопрос: а что, если мы СНАЧАЛА нашли $P_{1000}(0 \le m \le 105) \approx 0,7019$?
Тогда будет другая версия решения:
$P_{1000}(105 < m \le 1000) = 1 - P_{1000}(0 \le m \le 105) = 1 - 0,7019 = 0,2981$

Но как так может быть?! – в двух способах получаются разные ответы! Всё просто: интегральная теорема Лапласа – это
метод **приближённого** вычисления, и поэтому приемлемы оба пути.

Для более точных расчётов следует воспользоваться **формулой Бернулли** и, например, экселевской функцией *БИНОМРАСП*. В результате **её применения** получаем:
$P_{1000}^{0} + P_{1000}^{1} + P_{1000}^{2} + \ldots + P_{1000}^{104} + P_{1000}^{105} \approx 0,7221$

И я выражаю благодарность одному из посетителей сайта, который обратил внимание на эту тонкость – она выпала из моего поля зрения, так как исследование полной группы событий редко встречается на практике. Желающие могут ознакомиться с **содержательной дискуссией** по этому поводу.

Заключительный пример для самостоятельного решения:

Задача 6

В обычный учебный день вероятность присутствия студента на лекции равна 0,8. Найти вероятность того, что из 100 студентов на лекции будут присутствовать:

а) 85-90%;
б) половина студентов;
в) не менее 72 студентов.

Постарайтесь не пропускать задание ;-) Краткое решение и ответ совсем близко.

Здесь, несмотря на оговорку, все равно не всё гладко: известно, что процент прогулов у юношей заметно отличается от аналогичного показателя у девушек, поэтому усреднённая оценка несколько некорректна. Задачу следовало бы сформулировать для кадетского корпуса либо Института благородных девиц =) Неожиданно, но юноши, скорее всего, посещают занятия лучше =)

Вспомнилась, к слову, коварная задачка: вероятно ли встретить на улице 100 мужчин подряд? Запросто! Если навстречу прошагает рота солдат. Многие думают, что шансы встретить мужчину либо женщину составляют примерно 50 на 50 и даже встреча подряд десяти прохожих одного пола крайне маловероятна. Но почти все забывают об условии **равновозможности событий.** Так, например, если за углом находится отделение полиции или швейная фабрика, то встреча мужчины/женщины будет совсем не равновозможной.

Подобные моменты нужно обязательно учитывать в своих **статистических исследованиях**, которые бывают у каждого из нас хотя бы на бытовом уровне =)