1 Условная вероятность. Независимость

Первая опечатка не заставила долго себя ждать

Сегодня мы будем рассматривать пространства с, вообще говоря, бесконечным числом элементарных исходов, но не более чем счетным. При этом мы дополнительно требуем от вероятности выполнения свойства счетной аддитивности

$$\mathbf{P}(A_1 + \cdots + A_n + \cdots) = \sum_{n=1}^{\infty} \mathbf{P}(A_n).$$

Как и прежде мы будем задавать набор вероятностей p_i элементарных исходов, в сумме дающих единицу:

$$\sum_{i=1}^{\infty} p_i < +\infty$$

Отсюда автоматически задается вероятность любого события

$$\mathbf{P}(A) = \sum_{k:\ \omega_k \in A} \mathbf{P}(\omega_k) = \sum_{k:\ \omega_k \in A} p_k,$$

только теперь сумма может быть бесконечной.

1.1 Введение

Наше представление о вероятности события зависит от нашей информированности. Вероятность того, что у моего соперника среди пяти карт будет туз пик равна $5/52$. Если я знаю 5 карт в своей руке, то эта вероятность уже $5/47$ (или 0, конечно же). А если я помню три карты, которые обменял ходом раньше, то уже $5/44$. А если у него за спиной стоит зеркало и я вижу там туза пик, то она равна 1.

В наш век информации мы постоянно получаем новые сведения, которые в силу вышесказанного, меняют нашу вероятность. Сегодня мы поговорим о том, как работать в вероятностной модели в условиях поступающей информации.

1.1.1 Классический случай

Пусть $\Omega = \{\omega_1, \ldots, \omega_N\}$ — конечное пространство элементарных исходов с классически заданной вероятностью $\mathbf{P}(\omega_i) = 1/N$. Тогда вероятность события A определяется соотношением

$$\mathbf{P}(A) = \frac{|A|}{N}.$$

Предположим, что мы узнали о том, что выполнено событие B. Тогда B становится нашим новым вероятностным пространством, а наше событие A превращается в $A \cap B$. При этом из соображений симметрии все исходы остаются равновероятными, поэтому

$$\mathbf{P}_B(A) = \frac{|A \cap B|}{|B|}.$$

Можно заметить, что можно записать эту вероятность и в старых терминах.

$$\mathbf{P}_B(A) = \frac{|A \cap B|/N}{|B|/N} = \frac{\mathbf{P}(AB)}{\mathbf{P}(B)}.$$

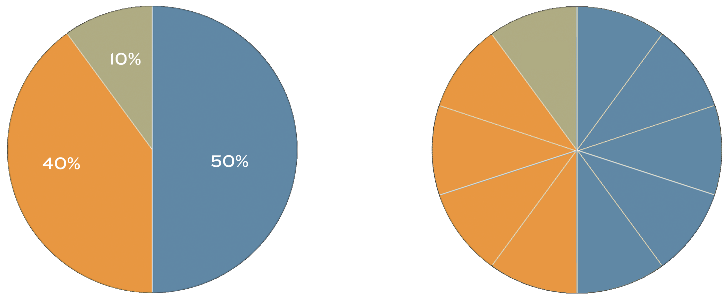
Итак, вероятность A "при условии, что случилось B" задается формулой $\mathbf{P}(AB)/\mathbf{P}(B)$.

1.1.2 "Рациональный" случай

Пусть $(\Omega, \mathcal{F}, \mathbf{P})$ — вероятностное пространство, где $\Omega = \{\omega_1, \ldots, \omega_N\}$ — конечное пространство элементарных исходов, вероятности исходов равны $p_1, \ldots, p_N$, соответственно, где $p_1 = a_1/b_1, \ldots, p_N = a_N/b_N$ — рациональные числа.

Введем новое пространство $\widetilde{\Omega} = \{\widetilde{\omega}_1, \ldots, \widetilde{\omega}_M\}$, где $M = b_1 \cdots b_N$, первые $a_1 b_2 \cdots .b_N$ новых исходов в объединении дают ω_1, вторые $b_1 a_2 b_3 \cdots b_N$ — ω_2 и так далее. Если мы определим вероятность $\widetilde{\mathbf{P}}$ на пространстве $\widetilde{\Omega}$ классическим образом, то $(\Omega, \mathcal{F}, \mathbf{P})$ и $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ окажутся согласованными — любом событие A имеет ту же вероятность $\mathbf{P}(A)$, что и соответствующее ему событие $\widetilde{A}$ из $\widetilde{\mathcal{F}}$. Иначе говоря, мы разбили каждый исход в старом

Рис. 1: Пример разбиения пространства с неклассически заданной вероятностью



пространстве Ω на части так, чтобы они получились равной вероятности.

Не так важно, как это реализовано физически, для нас существенно, что нашему пространству соответствует пространству с классически определенной вероятностью, на котором соображения симметрии и равноправия позволяют понять как надо определять вероятность. Мы будем прибегать к этому мысленному эксперименту и впредь — если мы хотим определить какое-то понятия с физически понятным подтекстом, то мы будем естественным образом определять его на классическом пространстве, а затем переносить его на произвольное.

В пространстве $\widetilde{\Omega}$ мы определяем вероятность события A "при условии, что случилось B" формулой $\widetilde{\mathbf{P}}(AB)/\widetilde{\mathbf{P}}(B)$. Значит, в старом пространстве естественно определять ее формулой

$$\frac{\mathbf{P}(AB)}{\mathbf{P}(B)}.$$

1.2 Условная вероятность

Где-то тут есть маленькая опечатка

Естественным образом, та же формула будет сохраняться и для иррациональных вероятностей исходов. Итак,

Определение 1. *Условной вероятностью* события A при условии B называют

$$\mathbf{P}(A|B) := \frac{\mathbf{P}(AB)}{\mathbf{P}(B)}.$$

Здесь предполагается, что $\mathbf{P}(B) \geq 0$.

Если рассуждения предыдущего раздела вам не понравились, то вы можете просто пропустить их и считать, что это определение и всё тут.

Содержательно мы понимаем под условной вероятностью вероятность события A, когда мы уже знаем, что случилось B.

Вопрос 1. При броске трех костей выпали три четных числа. Какая вероятность, что в сумме выпало 6 очков?
a) $1/27$;
b) $2/9$;
c) $1/81$;
d) $5/72$.

Пример 1. Пусть в урне 15 шаров из которых 5 белых и 10 черных шаров, мы выбираем три шара без возвращения. Пусть A — событие "первый шар белый", B — событие "второй шар белый", C — событие "третий шар черный".

Пространство элементарных исходов имеет вид

$$\Omega = \{(i,j,k), i \neq j, i \neq k, j \neq k, i,j,k \in \{1,\ldots,15\}\}, \quad A = \{(i,j,k) \in \Omega : i \in \{1,\ldots,5\}\},$$
$$B = \{(i,j,k) \in \Omega : j \in \{1,\ldots,5\}\}, \quad C = \{(i,j,k) \in \Omega : k \in \{6,\ldots,15\}\}.$$

Тогда
$$\mathbf{P}(A) = \frac{5 \cdot 14 \cdot 13}{15 \cdot 14 \cdot 13} = \frac{1}{3}, \quad \mathbf{P}(AB) = \frac{5 \cdot 4 \cdot 13}{15 \cdot 14 \cdot 13} = \frac{2}{21}, \quad \mathbf{P}(ABC) = \frac{5 \cdot 4 \cdot 10}{15 \cdot 14 \cdot 13} = \frac{10}{39}.$$
Следовательно,
$$\mathbf{P}(B|A) = \frac{5 \cdot 4 \cdot 13}{5 \cdot 14 \cdot 13} = \frac{4}{14}, \quad \mathbf{P}(C|AB) = \frac{5 \cdot 4 \cdot 10}{5 \cdot 4 \cdot 13} = \frac{10}{13}.$$
Ответы совпадают с нашими ожиданиями. Действительно, вероятность B при условии A — это вероятность вытянуть белый шар вторым из урны в мире, в котором первым вытянутым шар был белым. В этом мире в урне перед вторым вытягиванием было 14 шаров, из которых 4 белых. Точно также вероятность C при условии AB — это вероятность вытянуть черный шар из урны, где 10 черных и 3 белых шара.

Зачастую хочется записать что-то в духе двойной условной вероятности $\mathbf{P}((C|B)|A)$ (так писать неправильно!) — вероятность, что случится C, если уже случилось B, а перед этим случилось A. Однако, это просто $\mathbf{P}(C|AB)$ как из формального определения, так и из его физической интерпретации.

Будьте внимательны, недалеко целых две формулы с ошибками

Как мы видим из примера с шарами, зачастую мы легко можем найти условные вероятности (потому что легко работаем в новой реальности, заданной условием), а вот вероятности пересечения нам получить сложнее. В связи с этим полезно перевернуть определение условной вероятности:

$$\mathbf{P}(AB) = \mathbf{P}(A|B)\mathbf{P}(A).$$

Аналогичный результат можно доказать в случае n событий $A_1, \ldots, A_n$:

$$\mathbf{P}(A_n \cdots A_1) = \mathbf{P}(A_n|A_{n-1} \cdots A_1) \cdots \mathbf{P}(A_2|A_1)\mathbf{P}(A_1).$$

Этот результат называют теоремой умножения.

Доказательство. Докажем этот факт по индукции. При $n = 2$ он верен.

Пусть при $n = k$ он верен, докажем его при $n = k + 1$. В силу предположения и базы индукции

$$\mathbf{P}(A_{k+1} \cdots A_1) = \mathbf{P}(A_{k+1}|A_k)\mathbf{P}(A_k \cdots A_1) = \mathbf{P}(A_{k+1}|A_k)\mathbf{P}(A_k|A_{k-1}) \cdots \mathbf{P}(A_2|A_1)\mathbf{P}(A_1),$$

что и требовалось доказать. □

Эту формулу удобно использовать для подсчета вероятностей пересечения событий.

Вопрос 2. Чему равна вероятность того, что при трех вытягиваниях без возвращения из урны с 8 черными, 5 белыми и 3 красными шарами, мы вытянем черный, белый и снова черный шары.
a) 1/12
b) 5/64
c) 1/27
d) 4/21.

1.3 Формула полной вероятности и формула Байеса

в этом разделе тоже спрятана пара ошибок

Пример 2. Предположим, что мы хотим найти вероятность события B: второй шар, взятый из урны с 5 белыми и 10 черными шарами, будет белым.

Давайте разберемся какой шар выпал первым. Пусть A — событие, заключающееся в том, что первый шар белый. Мы легко можем найти

$$\mathbf{P}(AB) = \mathbf{P}(B|A)\mathbf{P}(A) = \frac{5}{15}\,\frac{4}{14} = \frac{2}{21}, \quad \mathbf{P}(\overline{A}B) = \mathbf{P}(B|\overline{A})\mathbf{P}(\overline{A}) = \frac{10}{15}\,\frac{5}{14} = \frac{5}{21},$$

откуда

$$\mathbf{P}(B) = \mathbf{P}(B|A)\mathbf{P}(A) + \mathbf{P}(B|\overline{A})\mathbf{P}(\overline{A}) = \frac{2}{21} + \frac{5}{21} = \frac{7}{21} = \frac{1}{3}.$$

Мы уже знаем этот ответ с прошлого занятия: вероятность, что второй шар белый та же, что и для первого. Но теперь мы получили новый инструмент получения такой формулы.

В общем случае аналогичная формула также справедлива. Сформулируем ее.

Определение 2. События $B_1, \ldots, B_N$ называются разбиением, если

$$\Omega = \bigcup_{n=1}^{N} B_n,$$

то есть если B_i не пересекаются, и их объединение есть все Ω. Мы предполагаем, что N может быть конечным или бесконечным.

Пусть $B_1, B_2, \ldots$. — события с ненулевой вероятностью, образующие разбиение Ω. Тогда

$$\mathbf{P}(A) = \mathbf{P}\left(\sum_{i=1}^{n} A \cup B_i\right) = \sum_{i=1}^{n} \mathbf{P}(AB_i) = \sum_{i=1}^{n} \mathbf{P}(A|B_i)\mathbf{P}(B_i).$$

Это соотношение называют *формулой полной вероятности*. Если число событий счетно, то формула остается в силе, но конечная сумма превращается в ряд.

Почему эта формула так полезна? Она позволяет посчитать вероятность в сложно устроенных экспериментах.

Пример 3. Пусть мы взяли то ли монету, на сторонах которой орел и решка, то ли монету, на сторонах которой два орла (равновероятно). Какая вероятность, что при трех бросаниях выпадет три орла?

Положим A – выпало 3 орла, B_1 – монета с орлом и решкой, B_2 – монета с двумя орлами. При этом

$$\mathbf{P}(A) = \mathbf{P}(A|B_1)\mathbf{P}(B_1) + \mathbf{P}(A|B_2)\mathbf{P}(A_2) = \frac{1}{8} \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{9}{16}.$$

Как мы видим, с помощью формулы полной вероятности можно свести сложную задачу к нескольким простым, не погружаясь в запутанное конструирование общего вероятностного пространства.

Вопрос 3. Мы подбрасываем монету с вероятностью орла 1/3. Если она выпадает на решку, то мы вытаскиваем шар из урны с 3 белыми и 7 черными шарами, а если на орла, то из урны с 9 белыми и 1 черным шаром. Какая вероятность того, что мы вытащим белый шар?
a) 3/5
b) 7/10.
c) 2/5.
d) 1/2.

Да и ниже не все гладко

Отметим еще одну формулу, легко вытекающую из формулы полной вероятности. Предположим, что мы знаем $\mathbf{P}(B_1), \ldots, \mathbf{P}(B_n)$ для разбиения $B_1, \ldots, B_n$. Пусть также известны $\mathbf{P}(A|B_1), \ldots, \mathbf{P}(A|B_n)$. Тогда

$$\mathbf{P}(B_1|A) = \frac{\mathbf{P}(A|B_1)\mathbf{P}(B_1)}{\mathbf{P}(A)} = \frac{\mathbf{P}(A|B_1)\mathbf{P}(B_1)}{\sum_{i=1}^{n} \mathbf{P}(B_i|A)\mathbf{P}(B_i)}.$$

Данная формула называется формулой Байеса.

Пример 4. Представим, что в последнем вопросе мы видим, что шар белый, и хотим восстановить на что упала монета. Тогда если B_1 — монета упала на орла, A — шар белый, то

$$\mathbf{P}(B_1|A) = \frac{\mathbf{P}(A|B_1)\mathbf{P}(B_1)}{\mathbf{P}(A|B_1)\mathbf{P}(B_1) + \mathbf{P}(A|\overline{B_1})\mathbf{P}(\overline{B_1})} = \frac{\frac{1}{3} \cdot \frac{9}{10}}{\frac{1}{3} \cdot \frac{9}{10} + \frac{2}{3} \cdot \frac{3}{10}} = \frac{3}{5}.$$

После вытягивания белого шара вероятность того, что монета упала на орла сильно возросла — от 0.(3) до 0.6, то есть почти вдвое.

Пример 5. Представим, что тест на определение болезни с вероятностью 99% определяет больным больного, с вероятностью 2% определяет здорового больным. При этом среди населения 1% больных людей. Человек сдал тест и тот определил его больным. Какая вероятность того, что он действительно болен?

Положим B_1 – человек болен, B_2 – здоров, A – тест признал человека больным. Заметим, что

$$\mathbf{P}(B_1) = 0.01, \quad \mathbf{P}(B_2) = 0.99, \quad \mathbf{P}(A|B_1) = 0.99, \quad \mathbf{P}(A|B_2) = 0.02.$$

Следовательно,

$$\mathbf{P}(B_1|A) = \frac{\mathbf{P}(A|B_1)\mathbf{P}(B_1)}{\mathbf{P}(A|B_1)\mathbf{P}(B_1) + \mathbf{P}(A|B_2)\mathbf{P}(B_2)} = \frac{0.010.99}{0.010.99 + 0.990.02} = \frac{1}{3}.$$

Несмотря на достаточную надежность теста, редкость болезни снижает шансы верного определения больных. Можно представить себе 100 случайных человек: из них всего 1 в среднем болен (и тест, вероятно, признает его больным), а 99 здоровы и тест в среднем 2 из них назовет больными.

1.4 Независимость

Это понятие является ключевым для развития теории вероятностей в отдельности от общей теории меры. Итак, дадим следующее естественное определение.

Определение 3. События A и B называются независимыми, если $\mathbf{P}(A|B) = \mathbf{P}(A)$ или $\mathbf{P}(B) = 0$.

Это определение в силу определения условной вероятности переписывается в виде

Определение 4. События A и B называются независимыми, если $\mathbf{P}(AB) = \mathbf{P}(A)\mathbf{P}(B)$.

Второе определение чаще всего удобнее для проверки и мы будем использовать его.

Начнем с нескольких простых замечаний, которые вам предстоит доказать самостоятельно:
Кроме одного, которое неверно

- Если A и B независимы, то A и $\overline{B}$ независимы, B и $\overline{A}$ независимы, $\overline{A}$ и $\overline{B}$ независимы
- События вероятности 0 и 1 независимы от любых событий.
- Событие независимо от самого себя только если оно имеет вероятность 0 или 1.
- Непересекающиеся события независимы только если одно из них имеет вероятность 0.
- Если A не зависит от B и A независит от C, где $BC = \emptyset$, то A независит от $B + C$.
- A не зависит от $\overline{A}$.

Чуть ниже что-то не так

Нам понадобится более общее определение независимости n событий.

Определение 5. События $A_1, \dots, A_n$ — независимы (в совокупности), если

$$\mathbf{P}(A_1 \cdots A_n) = \mathbf{P}(A_1) \cdots \mathbf{P}(A_n).$$

Давайте поймем почему нужно такое сложное определение. Неужели не хватит просто независимости каждого события с каждым (такую независимость называют попарной)?

Пример 6. Предположим, что мы бросаем симметричную монету дважды. Рассмотрим события A = {первый бросок оказался на орла}, B = {второй бросок оказался на орла}, C = {ровно один бросок оказался на орла}. Тогда

$$\mathbf{P}(A|B) = \mathbf{P}(A) = \frac{1}{2},\ \mathbf{P}(C|A) = \mathbf{P}(C) = \frac{1}{2},\ \mathbf{P}(C|B) = \mathbf{P}(C) = \frac{1}{2}.$$

Поэтому каждые два события независимы. Однако,

$$\mathbf{P}(ABC) = 0 \neq \mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C) = \frac{1}{8}.$$

Поэтому каждое из событий не влияет на шансы каждого из других случиться, но два события вместе делают невозможным третье.

Точно также можно построить n событий, любые $n-1$ из которых независимы, а все n вместе зависимы.

Даже не трогая доказательство – тут не все чисто

Можно привести следующий критерий независимости, обобщающий первое свойство независимости двух событий, указанное выше:

Лемма 1. *События $A_1, \ldots, A_n$ независимы тогда и только тогда, когда*

$$\mathbf{P}\left(A_1^{\delta_1} \cdots A_n^{\delta_n}\right) = \mathbf{P}\left(A_1^{\delta_1}\right) \cdots \mathbf{P}\left(A_n^{\delta_n}\right), \tag{1}$$

при всех $\delta_i \in \{0,1\}$, $A^1 = A$, $A^0 = 1$.

Доказательство. (Не входит в программу экзамена).

1) Докажем достаточность. Пусть $A_1, \ldots, A_n$ независимы. Покажем большее, чем нам требуется — что

$$\mathbf{P}(A_{i_1}^{\delta_{i_1}} \cdots A_{i_l}^{\delta_{i_l}}) = \mathbf{P}(A_{i_1}^{\delta_{i_1}}) \cdots \mathbf{P}(A_{i_l}^{\delta_{i_l}})$$

при любых l, $1 \le i_1 < \ldots < i_l \le n$, $\delta_{i_1}, \ldots, \delta_{i_l} \in \{0,1\}$. Доказательство проведем индукцией по $k = l - (\delta_{i_1} + \cdots . + \delta_{i_l})$.

База индукции: при $k = 0$ утверждение вытекает из того, что $A_1, \ldots, A_n$ независимы.

Переход индукции: пусть при $k \le m$ при некотором m утверждение доказано, докажем при $k = m+1$. Без ограничения общности считаем $\delta_{i_1} = 0$. Тогда

$$\mathbf{P}\left(A_{i_1}^0 A_{i_2}^{\delta_{i_2}} \cdots A_{i_l}^{\delta_{i_l}}\right) = \mathbf{P}\left(A_{i_2}^{\delta_{i_2}} \cdots A_{i_l}^{\delta_{i_l}}\right) - \mathbf{P}\left(A_{i_1}^1 A_{i_2}^{\delta_{i_2}} \cdots A_{i_l}^{\delta_{i_l}}\right) \tag{2}$$

при любых $i_1, \ldots, i_l$. В силу предположения индукции

$$\mathbf{P}\left(A_{i_2}^{\delta_{i_2}} \cdots A_{i_l}^{\delta_{i_l}}\right) = \mathbf{P}\left(A_{i_2}^{\delta_{i_2}}\right) \cdots \mathbf{P}\left(A_{i_l}^{\delta_{i_l}}\right),\ \mathbf{P}\left(A_{i_1}^1 A_{i_2}^{\delta_{i_2}} \cdots A_{i_l}^{\delta_{i_l}}\right) = \mathbf{P}\left(A_{i_1}^1\right) \mathbf{P}\left(A_{i_2}^{\delta_{i_2}}\right) \cdots \mathbf{P}\left(A_{i_l}^{\delta_{i_l}}\right). \tag{3}$$

Подставляя (3) в (2), получаем

$$\mathbf{P}\left(A_{i_1}^0 A_{i_2}^{\delta_{i_2}} \cdots A_{i_l}^{\delta_{i_l}}\right) = \left(1 - \mathbf{P}\left(A_{i_1}\right)\right) \mathbf{P}\left(A_{i_2}^{\delta_{i_2}}\right) \cdots \mathbf{P}\left(A_{i_l}^{\delta_{i_l}}\right) = \mathbf{P}\left(A_{i_1}^0\right) \mathbf{P}\left(A_{i_2}^{\delta_{i_2}}\right) \cdots \mathbf{P}\left(A_{i_l}^{\delta_{i_l}}\right).$$

Достаточность показана.

2) Докажем необходимость. Пусть мы знаем (1) для любого набора $i_1, \ldots, i_n$. Докажем, что

$$\mathbf{P}\left(A_{i_1} \cdots A_{i_l}\right) = \mathbf{P}\left(A_{i_1}\right) \cdots \mathbf{P}\left(A_{i_l}\right)$$

при всех $l \le n$. Для удобства будем считать, что $i_1 = 1, \ldots, i_l = l$ (это вопрос нумерации A_i). Тогда

$$\mathbf{P}\left(A_1 \cdots A_l\right) = \sum_{\delta_{l+1}, \ldots, \delta_n \in \{0,1\}} \mathbf{P}\left(A_1 \cdots A_l A_{l+1}^{\delta_{l+1}} \cdots A_n^{\delta_n}\right) = \sum_{\delta_{l+1}, \ldots, \delta_n \in \{0,1\}} \mathbf{P}\left(A_1\right) \cdots \mathbf{P}\left(A_l\right) \mathbf{P}\left(A_{l+1}^{\delta_{l+1}}\right) \cdots \mathbf{P}\left(A_n^{\delta_n}\right) =$$
$$\mathbf{P}\left(A_1\right) \cdots \mathbf{P}\left(A_l\right) \sum_{\delta_{l+1}, \ldots, \delta_n \in \{0,1\}} \mathbf{P}\left(A_{l+1}^{\delta_{l+1}}\right) \cdots \mathbf{P}\left(A_n^{\delta_n}\right).$$

Остается показать, что сумма в правой части равна единице. Для этого заметим, что это в точности

$$\left(\mathbf{P}\left(A_{l+1}^0\right) + \mathbf{P}\left(A_{l+1}^1\right)\right) \cdots \left(\mathbf{P}\left(A_n^0\right) + \mathbf{P}\left(A_n^1\right)\right) = 1 \cdot 1 \cdot \cdots \cdot 1 = 1.$$

Лемма доказана. □

Данная лемма удобна тем, что позволяет рассматривать только n событий, не рассматривая все поднаборы, что требует определение независимости.

1.5 Ответы на вопросы

1. Наше пространство состоит из пар $(2i, 2j)$, $i, j \in \{1, 2, 3\}$, следовательно в нем 9 исходов. Нам подходят 3 пары $(2, 6)$, $(4, 4)$, $(6, 2)$. Итого $1/3$.

2. В силу теоремы произведения
$$\frac{8}{16} \cdot \frac{5}{15} \cdot \frac{7}{14} = \frac{1}{12}.$$

3. В силу формулы полной вероятности искомая вероятность есть
$$\frac{2}{3} \cdot \frac{3}{10} + \frac{1}{3} \cdot \frac{9}{10} = \frac{5}{10} = \frac{1}{2}.$$