

Binomialverteilung

Когда есть фиксированное число попыток

Три раза повторяем опыт, в котором выбранное событие может произойти или не произойти. Нас интересует, случится ли оно ровно два раза. Это возможно в разном порядке: сначала два успеха, успехи с перерывом или два успеха в конце.

Поэтому одного произведения вероятностей недостаточно: нужно учесть все подходящие порядки. На этом небольшом примере удобно понять биномиальное распределение — **Binomialverteilung**. Оно описывает число успехов, когда попытки независимы и вероятность успеха остаётся одной и той же.

Почему нужно считать порядки

Обозначим успех буквой У, неуспех — Н. Ровно два успеха в трёх попытках дают порядки УУН, УНУ и НУУ. Если попытки независимы, а вероятность успеха каждый раз равна p , вероятность каждого порядка равна $p^2(1-p)$. Порядки несовместны, поэтому складываем их вероятности: получаем $3p^2(1-p)$.

При большем числе попыток перечислять порядки неудобно. Их число для k успехов из n равно $\binom{n}{k}$. Вероятность одного такого порядка состоит из k множителей p и $n-k$ множителей $1-p$. Поэтому

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k \in \{0, 1, \dots, n\}.$$

Здесь X — число успехов, $n \in \mathbb{N}$ — заранее фиксированное число попыток, $p \in [0, 1]$. Короткая запись этой модели: $X \sim B(n, p)$. Под успехом понимаем выбранный исход, например брак детали; это название не оценивает, хороший ли результат получился.

Запись $\binom{n}{k}$ читается «число сочетаний из n по k »: мы выбираем, на каких k местах стоят успехи. Например, три места из восьми можно по очереди выбрать $8 \cdot 7 \cdot 6$ способами. Но один набор мест при этом посчитан $3 \cdot 2 \cdot 1$ раз — во всех порядках выбора. Поэтому $\binom{8}{3} = (8 \cdot 7 \cdot 6) / (3 \cdot 2 \cdot 1) = 56$. В общем виде

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}, \quad n! = 1 \cdot 2 \cdots n, \quad 0! = 1.$$

При $k = 0$ или $k = n$ есть ровно один порядок: все неуспехи или все успехи.

В рассуждении использовались четыре условия: число попыток фиксировано, в каждой различаем успех и неуспех, вероятность p постоянна, попытки независимы. Без них произведение и подсчёт порядков могут дать неверную модель. На краях проще рассуждать напрямую: при $p = 0$ всегда $X = 0$, при $p = 1$ всегда $X = n$.

! Не биномиальная ситуация

Если выбор идёт **ohne Zurücklegen** и вероятность меняется после каждого выбора, это не обычное биномиальное распределение. Для точного расчёта часто нужно гипергеометрическое распределение (**hypergeometrische Verteilung**) или дерево. Биномиальную модель можно брать как приближение только если это разрешено и доля выборки мала.

Слова в знаки

M Формулировка EPM → запись

Фраза	Запись
genau 2	$P(X = 2)$
mindestens 2	$P(X \geq 2)$
mehr als 2	$P(X > 2) = P(X \geq 3)$
höchstens 2	$P(X \leq 2)$
weniger als 5	$P(X < 5) = P(X \leq 4)$
zwischen 2 und 5	если границы включены: $P(2 \leq X \leq 5)$

E Gegenwahrscheinlichkeit

Событию «хотя бы один успех» противоположно событие «ни одного». Вместе они исчерпывают все варианты, поэтому их вероятности дают 1:

$$P(X \geq 1) = 1 - P(X = 0).$$

Если требуется хотя бы два успеха, убираем уже два неподходящих количества: ноль и один.

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1).$$

Суммы и интервалы

От одного значения к нескольким

Число успехов не может одновременно равняться нулю, одному и двум. Чтобы найти вероятность «не более k », складываем вероятности всех допустимых количеств от 0 до k . Знак $\sum$ сокращает именно эту сумму; буква i по очереди принимает значения $0, 1, \dots, k$:

$$P(X \leq k) = \sum_{i=0}^k \binom{n}{i} p^i (1-p)^{n-i}.$$

Для целого интервала $0 \leq a \leq b \leq n$:

$$P(a \leq X \leq b) = P(X \leq b) - P(X \leq a-1).$$

Частая ошибка.

$a-1$ появляется только потому, что X принимает целые значения: $0, 1, 2, \dots, n$. Это не формула для нормального распределения.

ex Коротко

$$P(5 \leq X \leq 9) = P(X \leq 9) - P(X \leq 4).$$

Мы вычитаем всё до 4, чтобы оставить 5, 6, 7, 8, 9.

v Визуализация интервала для $P(5 \leq X \leq 9)$

Считаем целые значения X от 0 до 10:



3 типовых задания

Возьмём 8 независимых попыток с вероятностью успеха 0,25 в каждой. Число успехов $X \sim B(8, 0,25)$ остаётся тем же во всех трёх вопросах; меняется только набор подходящих значений X .

ex 1. genau

$$n = 8, \quad p = 0,25, \quad \text{genau 3 Erfolge} \Rightarrow P(X = 3).$$

$$P(X = 3) = \binom{8}{3} (0,25)^3 (0,75)^5.$$

Здесь 56 порядков, в каждом три успеха и пять неудач. Получаем $P(X = 3) \approx 0,2076$, то есть примерно 20,76%.

ex 2. höchstens

$$\text{höchstens 3 Erfolge} \Rightarrow P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3).$$

На калькуляторе это накопленная вероятность до 3 включительно.

ex 3. Ловушка со словом mindestens

$$\text{mindestens 2 Erfolge} \Rightarrow P(X \geq 2) = 1 - P(X = 0) - P(X = 1).$$

Неправильно: считать только $P(X = 2)$.

Разбор задачи с mindestens 2

Streaming: задача 6 из снимка

Выбирают 10 клиентов. Нужно найти вероятность, что не менее двух (**mindestens 2**) из них старше 40 лет.

Erfolg:

älter als 40 Jahre.

По условию 70% клиентов не старше 40 лет. Остальные старше 40, поэтому вероятность выбранного нами успеха равна $1 - 0,70$:

$$p = 0,30.$$

Для биномиальной модели предполагаем независимость выборов и неизменное p .

$$X = \text{Anzahl der Kunden älter als 40}, \quad X \sim B(10, 0,30).$$

Проще исключить ноль и одного клиента старше 40, чем складывать вероятности всех количеств от 2 до 10:

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1).$$

$$P(X \geq 2) = 1 - 0,7^{10} - \binom{10}{1} 0,3 \cdot 0,7^9 \approx 0,8507.$$

Die Wahrscheinlichkeit beträgt etwa 85,1%.

Минимальное число попыток n

Нужно найти минимальное число попыток n , чтобы с вероятностью не меньше P_0 (например, 95%) получить хотя бы один успех (**mindestens ein Erfolg**). Далее предполагаем $0 < p < 1$ и $0 < P_0 < 1$.

Формулировка события:

$$P(X \geq 1) \geq P_0.$$

Переходим к противоположному событию (0 успехов):

$$1 - P(X = 0) \geq P_0 \quad \Rightarrow \quad 1 - (1 - p)^n \geq P_0.$$

Решаем неравенство относительно n :

$$(1 - p)^n \leq 1 - P_0.$$

Обе части положительны. Логарифм возрастает, поэтому при его применении знак неравенства сохраняется; показатель n становится множителем:

$$\ln((1 - p)^n) \leq \ln(1 - P_0) \quad \Rightarrow \quad n \cdot \ln(1 - p) \leq \ln(1 - P_0).$$

! Важный шаг: разворот знака

Так как $p \in (0, 1)$, то $(1 - p) < 1$, а значит $\ln(1 - p) < 0$. При делении на отрицательное число **знак неравенства переворачивается**:

$$n \geq \frac{\ln(1 - P_0)}{\ln(1 - p)}.$$

Полученное число n округляем **вверх** до ближайшего целого. Затем проверяем минимальность: найденное n должно подходить, а $n - 1$ — ещё нет.

ex Пример: хотя бы одна шестёрка

Сколько раз нужно бросить кубик, чтобы хотя бы одна шестёрка выпала с вероятностью не менее 95%?

$$p = \frac{1}{6}, \quad P_0 = 0,95 \quad \Rightarrow \quad 1 - p = \frac{5}{6}, \quad 1 - P_0 = 0,05.$$

$$P(X \geq 1) \geq 0,95 \quad \Rightarrow \quad 1 - \left(\frac{5}{6}\right)^n \geq 0,95 \quad \Rightarrow \quad \left(\frac{5}{6}\right)^n \leq 0,05.$$

$$n \geq \frac{\ln(0,05)}{\ln(5/6)} \approx \frac{-2,9957}{-0,1823} \approx 16,43 \quad \Rightarrow \quad n = 17.$$

$$1 - \left(\frac{5}{6}\right)^{16} \approx 0,9459 < 0,95, \quad 1 - \left(\frac{5}{6}\right)^{17} \approx 0,9549 \geq 0,95.$$

Ответ: Нужно бросить кубик не менее 17 раз.

Большое n : нормальное приближение

250 клиентов

Пусть

$X = \text{Anzahl der Kunden höchstens 40 Jahre.}$

Тогда:

$$X \sim B(250, 0,70).$$

Точная модель остаётся биномиальной. Её параметры:

$$E(X) = np, \quad \text{Var}(X) = np(1 - p), \quad \sigma = \sqrt{np(1 - p)}.$$

$E(X)$, или математическое ожидание (**Erwartungswert**), — среднее число успехов при многократном повторении всей серии. Оно может быть нецелым; отдельная серия не обязана давать ровно np успехов. Дисперсия $\text{Var}(X)$ — среднее значение квадрата отклонения от этого центра. Её корень σ возвращает масштаб разброса к единицам числа клиентов.

Точную сумму можно получить функцией биномиального распределения на калькуляторе. Для нормального приближения в этой серии используем практическое правило:

$$np = 175 \geq 10, \quad n(1-p) = 75 \geq 10.$$

Оба числа достаточно велики, поэтому биномиальные вероятности приближаем нормальной моделью

$$X_N \sim N\left(\mu = 175, \sigma = \sqrt{52,5}\right), \quad \sigma \approx 7,246.$$

weniger als 170:

$$X < 170 \Rightarrow X \leq 169.$$

Из-за перехода от целых значений к непрерывной кривой ставим границу посередине между 169 и 170:

$$P(X < 170) = P(X \leq 169) \approx P(X_N \leq 169,5).$$

Для нормальной кривой вероятность — площадь слева от границы. Вычтем центр 175 и разделим расстояние на σ ; получим стандартную координату z . Функция $\Phi(z)$ даёт левую площадь для нормальной кривой с центром 0 и стандартным отклонением 1:

$$z = \frac{169,5 - 175}{\sqrt{52,5}} \approx -0,759, \quad P \approx \Phi(-0,759) \approx 0,2239.$$

Die Wahrscheinlichkeit beträgt näherungsweise 22,4%.

Abweichung vom Erwartungswert

$$\mu = E(X) = 250 \cdot 0,70 = 175, \quad \sigma = \sqrt{52,5}.$$

$$\text{um maximal 10 abweichen} \Rightarrow |X - 175| \leq 10.$$

Модуль измеряет расстояние до 175 без учёта направления. Отойти не более чем на 10 можно до $175 - 10$ влево и до $175 + 10$ вправо:

$$165 \leq X \leq 185.$$

С поправкой на непрерывность интервал расширяется на 0,5 с обеих сторон:

$$P(165 \leq X \leq 185) \approx P(164,5 \leq X_N \leq 185,5).$$

$$z_1 \approx -1,449, \quad z_2 \approx 1,449.$$

Вычитаем из площади слева от правой границы площадь слева от левой:

$$P \approx \Phi(1,449) - \Phi(-1,449) \approx 0,8527.$$

Die Wahrscheinlichkeit beträgt näherungsweise 85,3%.

Е Приближение и точный расчёт

Для этих данных точный биномиальный расчёт даёт $P(X < 170) \approx 22,28\%$, а $P(165 \leq X \leq 185) \approx 85,30\%$. Нормальное приближение с поправкой даёт соответственно 22,39% и 85,27%. Выбор приближения и округление должны быть видны в записи. Правило $np, n(1-p) \geq 10$ не гарантирует заданной точности в хвостах.

М Как писать на экзамене

Строка решения	Что написать
Zufallsvariable	$X = \text{Anzahl der Erfolge}$
Modell	$X \sim B(n, p)$
Ereignis	например $P(X \geq 2)$, не словами
Вычисление	формула, противоположное событие или результат калькулятора
Ответ	вероятность в процентах с округлением

Типичные ошибки

! Что проверить в решении

- не записали $X = \text{Anzahl der Erfolge}$;
- не записали $X \sim B(n, p)$;
- **mindestens 2** посчитали как $P(X = 2)$;
- **weniger als 170** посчитали как $P(X \leq 170)$;
- в интервале забыли, что X целое;
- при большом n применили нормальное приближение без проверки np и $n(1-p)$ или забыли поправку на непрерывность 0,5;
- перепутали успех и противоположное событие;
- в задаче на минимальное n забыли перевернуть знак неравенства при делении на отрицательный $\ln(1-p)$.

Итоговая памятка

$$X \sim B(n, p), \quad P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

$$E(X) = n \cdot p, \quad \sigma = \sqrt{n \cdot p \cdot (1-p)}.$$

$$P(a \leq X \leq b) = P(X \leq b) - P(X \leq a-1).$$

$$np, n(1-p) \geq 10: \quad X_N \sim N\left(np, \sqrt{np(1-p)}\right), \quad [a; b] \mapsto [a - 0,5; b + 0,5].$$

Подсказки для калькулятора (GTR/CAS)

Названия, порядок аргументов и синтаксис зависят от модели калькулятора. Часто встречаются функции:

- **binomPdf(n, p, k)** — для точной вероятности $P(X = k)$;
- **binomCdf(n, p, k)** — для накопленной вероятности $P(X \leq k)$ от 0 до k ;
- для интервала всегда надёжно работает разность $P(X \leq b) - P(X \leq a-1)$; четырёхаргументный вызов есть не на всех устройствах.
- при обоснованном нормальном приближении используем **normCdf** с $\mu = np$, $\sigma = \sqrt{np(1-p)}$ и уже исправленными границами, например 169,5 вместо 169.

Перед экзаменом один раз сверь ввод с меню своего устройства.

Проверка результата: вероятность лежит в $[0; 1]$, успех определён словами, а в ответе указаны проценты и округление.

Самопроверка: сначала без ответа**Übungsaufgabe: Treffer**

Bei jedem Versuch beträgt die Trefferwahrscheinlichkeit $p = 0,2$. Die Versuche sind unabhängig.

1. Definieren Sie für 12 Versuche eine Zufallsvariable und geben Sie die Verteilung an.
2. Berechnen Sie die Wahrscheinlichkeit für genau 2 Treffer und für mindestens 2 Treffer.
3. Wie viele Versuche sind mindestens nötig, damit die Wahrscheinlichkeit für mindestens einen Treffer wenigstens 95% beträgt?

В последнем пункте проверь и найденное n , и $n - 1$.

Ответы к самопроверке

Lösung

X = Anzahl der Treffer in 12 Versuchen, $X \sim B(12, 0,2)$.

$$P(X = 2) = \binom{12}{2} 0,2^2 0,8^{10} \approx 0,2835.$$

$$P(X \geq 2) = 1 - 0,8^{12} - 12 \cdot 0,2 \cdot 0,8^{11} \approx 0,7251.$$

Für eine variable Anzahl n gilt:

$$1 - 0,8^n \geq 0,95 \quad \Rightarrow \quad n \geq \frac{\ln 0,05}{\ln 0,8} \approx 13,43.$$

Also sind mindestens 14 Versuche nötig:

$$1 - 0,8^{13} \approx 0,9450 < 0,95, \quad 1 - 0,8^{14} \approx 0,9560 \geq 0,95.$$

Ответы: примерно 28,35%, 72,51% и 14 испытаний. Проверка модели нужна до формулы: одинакового p без независимости недостаточно.

Содержание

Нажми название или номер страницы. Закладки PDF повторяют разделы. Ответы открывай после собственной попытки.

Почему нужно считать порядки	1	Минимальное число попыток n	5
Слова в знаки	2	Большое n : нормальное приближение	6
Суммы и интервалы	3	Типичные ошибки	8
3 типовых задания	4	Самопроверка: сначала без ответа	9
Разбор задачи с mindestens 2	5	Ответы к самопроверке	10