

Санкт-Петербургский государственный университет

Прикладная математика и информатика

Прикладная кибернетика

Научная исследовательская работа

СРАВНЕНИЕ НА СИНТЕТИЧЕСКИХ ДАННЫХ И НА ДАННЫХ С KAGGLE  
МЕТОДОВ ДЛЯ РАССЧЁТА ДОВЕРИТЕЛЬНЫХ ИНТЕРВАЛОВ ДЛЯ 90, 99  
КВАНТИЛЕЙ

Выполнил:

Кривоносов Тимофей Игоревич

Научный руководитель:

д. ф.-м. н., профессор

М. В. Юлдашев

Санкт-Петербург

2023

# Оглавление

1.	Введение . . . . .	3
2.	Методы нахождения доверительных интервалов . . . . .	5
2.1.	Наивный подход . . . . .	5
2.2.	Бустерэп . . . . .	7
2.3.	Дельта подход . . . . .	8
3.	Сбор и генерация данных . . . . .	11
3.1.	Синтезированные данные . . . . .	11
3.2.	Данные из Kaggle . . . . .	11
4.	Анализ и сравнение разобранных методов . . . . .	13
4.1.	Применение методов . . . . .	13
4.2.	Анализ и сравнение методов . . . . .	13
4.3.	Подведение итогов . . . . .	13
5.	Заключение . . . . .	14
6.	Приложение . . . . .	15

## 1. Введение

В современном мире статистический анализ играет важную роль во многих областях, от науки до бизнеса. При работе с данными мы часто сталкиваемся с необходимостью оценки параметров и построения доверительных интервалов для различных квантилей распределений. Квантили представляют собой значения, разделяющие вероятностное распределение на равные части.

В данной курсовой работе мы сосредоточимся на сравнении методов для расчета доверительных интервалов для 90-го и 99-го квантилей на основе синтетических данных и данных, предоставленных платформой Kaggle. Синтетические данные являются созданными искусственным образом наборами данных, которые позволяют нам контролировать различные аспекты, такие как распределение, объем выборки и т.д. Данные, полученные с Kaggle, представляют собой реальные данные, предоставленные сообществом пользователей этой платформы.

Для синтеза данных и проверки методов поиска доверительных интервалов для 90 и 99 квантилей в задаче сравнения производительности и задержки (latency), проведем следующие шаги:

- **Собрать данные:** Запустите эксперименты или тесты, которые измеряют производительность и задержку вашей системы. Запишите результаты для каждого теста.
- **Подготовить выборки:** Создайте выборки из результатов тестов, соответствующие каждому методу или условию, которое вы хотите сравнить. Например, если вы сравниваете два разных алгоритма, создайте две выборки, содержащие результаты для каждого алгоритма.
- **Анализировать выборки:** Используйте статистические методы для анализа выборок и вычисления доверительных интервалов. Для нахождения доверительного интервала 90% квантили, найдите 5-й и 95-й перцентили выборки. Для доверительного интервала 99% квантили найдите 0.5-й и 99.5-й перцентили выборки.
- **Проверьте методы поиска доверительных интервалов:** Сравните результаты, полученные различными методами поиска доверительных интервалов. Убедитесь, что методы возвращают адекватные и интерпретируемые результаты.

- Интерпретация результатов: Проанализируйте полученные доверительные интервалы и сделайте выводы относительно производительности и задержки системы. Например, если интервалы значительно отличаются для двух методов, это может указывать на значимые различия между ними.

Основной целью нашего исследования является провести сравнительный анализ методов расчета доверительных интервалов для выбранных квантилей на основе обоих типов данных. Это позволит нам оценить, насколько точно и надежно эти методы работают в различных условиях и с разными типами данных.

Ожидается, что результаты данного исследования помогут нам лучше понять применимость и эффективность различных методов для расчета доверительных интервалов на основе разных типов данных, что может быть полезным при принятии статистических решений в реальных ситуациях.

## 2. Методы нахождения доверительных интервалов

Доверительные интервалы являются важным инструментом в статистике и анализе данных. Они предоставляют информацию о диапазоне значений, в пределах которого с определенной вероятностью находится истинное значение параметра популяции. Методы нахождения доверительных интервалов разрабатываются для различных статистических оценок, таких как среднее, медиана, квантили и дисперсия.

В данном контексте существует несколько методов, включая аналитические, бутстреп и другие. Аналитические методы обычно требуют явных аналитических моделей и предположений о распределении данных. Бутстреп - это метод ресемплинга, который позволяет оценивать доверительные интервалы на основе эмпирических данных без явных предположений.

В этом контексте, мы будем исследовать различные методы нахождения доверительных интервалов и сравнивать их производительность в оценке статистических параметров в задачах анализа производительности и задержки.

### 2.1. Наивный подход

Наивный подход (или иногда "классический подход") при поиске доверительного интервала предполагает использование простых формул или методов для оценки параметров популяции или статистических характеристик на основе имеющихся данных без учета сложных статистических моделей или методов ресемплинга.

Основные черты наивного подхода:

- Простота: Наивный подход обычно предполагает использование простых статистических оценок, таких как выборочное среднее, выборочная дисперсия, медиана и квантили, без необходимости в более сложных методах.
- Явные формулы: В большинстве случаев наивные методы имеют явные аналитические формулы для вычисления оценок и доверительных интервалов. Например, доверительный интервал для среднего значения в выборке с известной дисперсией может быть вычислен с использованием формулы Z-критерия.
- Предположения о распределении: Наивные методы часто основываются на предположениях о распределении данных, таких как нормальное распределение. Они могут быть неустойчивыми к нарушениям этих предположений.

- Ограниченная гибкость: Наивные методы могут быть ограничены в способности учитывать сложные зависимости в данных или рассматривать данные, не подчиняющиеся стандартным распределениям.

Наивный подход полезен, когда у вас есть хорошие основания для использования простых методов, и когда они дают приемлемые результаты. Однако в случае сложных данных или данных, не соответствующих предполагаемым распределениям, наивный подход может быть недостаточным и неустойчивым. В таких случаях более сложные методы, такие как бутстреп или аналитические модели, могут быть более надежными.

Алгоритм нахождения доверительного интервала с его помощью обычно включает следующие шаги:

- Шаг 1: Собрать выборку данных. Это могут быть измерения, результаты эксперимента или другие данные, на основе которых вы хотите построить доверительный интервал.
- Шаг 2: Выбрать уровень доверия (например, 95 % или 99%). Уровень доверия определяет, с какой вероятностью истинное значение параметра популяции находится внутри доверительного интервала.
- Шаг 3: Определить статистическую оценку параметра. Наивный подход часто использует выборочные статистики, такие как выборочное среднее  $\bar{x}$  или выборочная медиана, для оценки параметра популяции.
- Шаг 4: Определить стандартную ошибку. Стандартная ошибка (SE) измеряет разброс выборочной статистики и зависит от объема выборки. Для среднего значения это может быть вычислено как стандартное отклонение выборки, деленное на корень из размера выборки ( $SE = \sigma / \sqrt{n}$ , где  $\sigma$  - стандартное отклонение популяции).
- Шаг 5: Определить критическое значение. Критическое значение определяет, какие значения статистики считаются "достаточно экстремальными" для данного уровня доверия. Наивный подход обычно использует критические значения из стандартных распределений, таких как Z-распределение или t-распределение.
- Шаг 6: Вычислить доверительный интервал. Доверительный интервал определяется как оценка параметра плюс/минус критическое значение, умноженное на

стандартную ошибку. Для среднего значения (наивный метод):

Доверительный интервал = оценка параметра  $\pm$  (критическое значение \* SE)

Пример:

Предположим, у нас есть выборка из 30 измерений длины стержня, и мы хотим построить 95 % доверительный интервал для средней длины стержня. Стандартное отклонение выборки равно 2 см.

- Выборка данных: [10, 11, 12, 14, 15, 11, 13, 12, 14, 15, 15, 13, 12, 13, 14, 16, 17, 18, 15, 12, 13, 12, 14, 15, 16, 17, 16, 13, 12, 14]
- Уровень доверия: 95%
- Статистическая оценка: Выборочное среднее  $\bar{x} = 14$
- Стандартная ошибка:  $SE = 2/\sqrt{30} \approx 0.365$
- Критическое значение: Для 95% доверия с использованием Z-распределения, критическое значение равно примерно 1.96.
- Вычисление доверительного интервала:

Доверительный интервал =  $14 \pm (1.96 * 0.365) \approx [13.29, 14.71]$

Таким образом, с 95% уровнем доверия мы можем утверждать, что средняя длина стержня находится в интервале от 13.29 см до 14.71 см.

## 2.2. Бустерэп

Бутстрэп - это метод статистической ресемплинга, который используется для оценки дисперсии и доверительных интервалов для статистических оценок, когда точное аналитическое решение сложно или невозможно. Он широко применяется в статистике и машинном обучении для решения различных задач, таких как оценка параметров, построение доверительных интервалов и проверка гипотез. Вот как он работает:

- Создание репликаций: Из вашей исходной выборки размером N (обычно с заменой) генерируются множество новых выборок (репликаций). Количество репликаций, обычно обозначаемое как B, может быть выбрано пользователем и обычно составляет несколько сотен или тысяч.

- Оценка интересующей статистики: Для каждой из  $B$  репликаций, вы оцениваете интересующую вас статистику. Это может быть, например, среднее, медиана, доля, стандартное отклонение, коэффициент корреляции или что-либо еще в зависимости от вашей задачи.
- Анализ распределения: После оценки статистики для каждой репликации, вы получаете распределение значений этой статистики на основе бутстреп-выборок.
- Построение доверительных интервалов: Из анализа этого распределения можно построить доверительные интервалы для вашей статистики. Наиболее распространенным методом для этого является метод "перцентильного доверительного интервала" который использует квантили распределения.

Пример использования метода:

Предположим, у нас есть выборка размером 100 наблюдений и мы хотим построить доверительный интервал для среднего значения. Мы можем применить метод бутстрэп следующим образом:

- Создаем множество бутстрэп-выборок путем случайного выбора 100 наблюдений из исходной выборки с возвращением.
- Для каждой бутстрэп-выборки вычисляем среднее значение.
- Полученные средние значения образуют распределение, которое отражает неопределенность вокруг истинного среднего значения.
- Используем полученное распределение для построения доверительного интервала, например, 95% доверительного интервала будет содержать средние значения, лежащие между 2.5-м и 97.5-м перцентилями этого распределения.

Таким образом, метод бутстрэп позволяет оценить статистическую неопределенность и построить доверительные интервалы для различных параметров на основе анализа повторных выборок из исходной выборки.

### 2.3. Дельта подход

Дельта-метод - это статистический метод, используемый для оценки дисперсии и построения доверительных интервалов для функций случайных величин на основе



их оценок и производных. Основное предположение метода Дельта состоит в том, что оценки параметров близки к нормальному распределению.

Доказательство метода Дельта является более сложным и требует математической формализации и использования разложения в ряд Тейлора. Давайте представим общий контур доказательства метода Дельта, который включает следующие шаги:

Предположим, у нас есть функция случайной величины  $X$ , которую мы обозначим как  $g(X)$ , и мы хотим оценить ожидание этой функции, то есть  $E[g(X)]$ .

- Разложение в ряд Тейлора: Начнем с разложения в ряд Тейлора для функции  $g(X)$  вокруг некоторого фиксированного значения  $a$ . Это разложение будет иметь вид:

$g(X) = g(a) + g'(a)(X - a) + g''(a)(X - a)^2/2 + \dots$  где  $g'(a)$  представляет производную функции  $g(X)$  в точке  $a$ .

- Взятие ожидания: Теперь возьмем ожидание от обеих сторон уравнения:

$$E[g(X)] = E[g(a) + g'(a)(X - a) + g''(a)(X - a)^2/2 + \dots]$$

- Использование линейности ожидания: Ожидание линейно, поэтому мы можем разбить его на отдельные члены:

$$E[g(X)] = g(a) + E[g'(a)(X - a)] + E[g''(a)(X - a)^2/2] + \dots$$

- Оценка среднего и дисперсии: Мы видим, что первый член  $g(a)$  - это константа, и остальные члены представляют собой смешанные моменты случайной величины  $X$ . Мы можем использовать оценки математического ожидания и дисперсии для оценки  $E[g(X)]$ :

$$E[g(X)] \approx g(a) + g'(a)E(X - a) + (1/2)g''(a)E((X - a)^2)$$

- Замена параметров: Теперь мы заменяем параметры  $a$ ,  $g(a)$ ,  $g'(a)$  и  $g''(a)$  на их оценки, соответственно. Обозначим оценки как  $\hat{a}$ ,  $g(\hat{a})$ ,  $g'(\hat{a})$  и  $g''(\hat{a})$ :

$$E[g(X)] \approx g(\hat{a}) + g'(\hat{a})E(X - \hat{a}) + (1/2)g''(\hat{a})E((X - \hat{a})^2)$$

- Оценка дисперсии: Мы можем оценить дисперсию случайной величины  $X - \hat{a}$  и заменить  $E((X - \hat{a})^2)$  на  $Var(X - \hat{a})$ :

$$E[g(X)] \approx g(\hat{a}) + g'(\hat{a})E(X - \hat{a}) + (1/2)g''(\hat{a})Var(X - \hat{a})$$

- Итоговый результат: Таким образом, мы получаем оценку  $E[g(X)]$  в зависимости от оценок параметров функции  $g(X)$ , а также смешанных моментов и дисперсии случайной величины  $X - \hat{a}$ . Это представляет собой общий контур доказательства метода Дельта, который объясняет, как мы можем оценить математическое ожидание функции случайной величины, используя разложение в ряд Тейлора и оценку параметров функции и их производных.

Чтобы использовать метод Дельта для квантильных метрик (например, для оценки доверительного интервала квантиля), выполните следующие шаги:

- Оцените квантиль на основе вашей выборки. Обозначьте это значение как оценку квантиля (например,  $\hat{q}$ ).
- Вычислите производную функции квантиля по параметру распределения (обычно параметр  $p$ ). Это даст вам информацию о градиенте функции вблизи оценки квантиля.
- Вычислите значение производной в точке, соответствующей вашей оценке квантиля (например,  $\hat{q}(p)$ ). Это будет вашей оценкой производной.
- Вычислите дисперсию оценки квантиля,  $Var(\hat{q})$ .
- Теперь, используя метод Дельта, можно вычислить доверительный интервал для оценки квантиля:
- Нижняя граница интервала:  $\hat{q} - Z * \sqrt{Var(\hat{q})}$   
Верхняя граница интервала:  $\hat{q} + Z * \sqrt{Var(\hat{q})}$

Здесь  $Z$  - это критическое значение для выбранного уровня доверия (например, для 95% уровня доверия,  $Z \approx 1.96$ ).

*Более углубленно дельта метод и пример его использования представлен в приложении 2 [2]*

Метод Дельта позволяет оценивать дисперсию и строить доверительные интервалы для различных функций от случайных величин, включая квантили. Это полезный метод в статистике и анализе данных, который позволяет оценить не только средние значения, но и другие статистики и метрики с учетом их дисперсии и нормальности распределения.

### 3. Сбор и генерация данных

#### 3.1. Синтезированные данные

При создании тестовых данных для проверки методов поиска доверительного интервала следует учитывать следующие аспекты:

- **Известное распределение данных:** Выбрать распределение, которое соответствует данным или сценарию исследования. Например, нормальное распределение, равномерное распределение или экспоненциальное распределение. (В нашем случае необходимо проверить все)
- **Размер выборки:** Определить размер выборки, который будет использоваться для генерации данных. (1000 эл)
- **Генерация случайных чисел:** Использовать подходящий генератор случайных чисел, чтобы создать данные, соответствующие выбранному распределению и параметрам.
- **Повторяемость** Убедится, что тестовые данные повторяемы, то есть при каждом запуске теста они будут генерироваться с одинаковыми значениями.

Во время создания программы для генерации данных необходимо учесть, что у каждого метода есть хорошие стороны и обратные, поэтому при синтезе нужно готовить выборки как одного вида, так и другого. Для этого будем использовать разные распределения: нормальное, равномерное и экспоненциальное. После проверим каждый из методов и проведем анализ. Программа для создания выборок находится в приложении 1 [1] или на GitHub.

#### 3.2. Данные из Kaggle

Также следует проверить методы на данных созданных сообществом программистов и аналитиков. Для этого зарегистрируемся на ресурсе Kaggle (Kaggle - это платформа для соревнований по анализу данных и машинному обучению. Она предоставляет сообществу специалистов по данным доступ к наборам данных, инструменты для исследования и моделирования данных, а также возможность участвовать в соревнованиях, где участники решают реальные проблемы с использованием машинного обучения и

статистического анализа. Kaggle также предлагает образовательные материалы, форумы и средства для сотрудничества над проектами.)

## 4. Анализ и сравнение разобранных методов

### 4.1. Применение методов

Создадим три программы, каждая из которых на вход получает файл с выборками, находит доверительный интервал, время выполнения метода и False Positive Rate.

Выходными данными является файл ".txt" имеющий вид

*(№) 1 (Доверительный интервал) [-1.599138 ; 1.65090] (Время поиска интервала) 6.91 (FPR) 0.098*

Программы и выходные данные находятся в приложении или на GitHub.

### 4.2. Анализ и сравнение методов

Получив результаты от каждого метода сравним их (программа - Приложение 3). Самым быстрым методом является (...) со средней скоростью: В свою очередь самым

### 4.3. Подведение итогов

## 5. Заключение

more...

## 6. Приложение

### 1. Программа генерации выборок:

```

DataFactory.py > ...
1  import numpy as np
2  import os
3
4  # Создание директории для сохранения файлов
5  save_path = "Users/krivosovti/Курсач/DataExponential"
6
7  # Создание директории для сохранения файлов
8  if not os.path.exists(save_path):
9      os.makedirs(save_path)
10
11  num_samples = 150
12  lambd = 1
13  # Генерация и сохранение выборок
14  for i in range(num_samples):
15
16      # Генерация случайной выборки
17      #sample = np.random.normal(loc=0, scale=1, size=1000) # Пример: нормальное распределение
18      #sample = np.random.uniform(size=1000) # Пример: равномерное распределение
19      sample = np.random.exponential(scale=1/lambd, size=1000) # Пример: экспоненциальное распределение
20      lambd = lambd / 1.4
21      # Создание пути и имени файла
22      file_name = f"sampleExponential_{i+1}.txt"
23      file_path = os.path.join(save_path, file_name)
24
25      # Сохранение выборки в файл
26      with open(file_path, "w") as file:
27          for value in sample:
28              file.write(f"{value}\n")

```

### 2. Дельта метод углубленно Методология: Предположим, что А/В-тест задается с несколькими вариантами, где кандидаты в каждом варианте получают разный опыт. Мы заинтересованы в измерении того, как

Опыт в каждом варианте влияет на  $q$ -й квантиль времени загрузки страницы. Чтобы измерить это влияние и вычислить статистическую значимость, нам нужны оценки количества выборки и стандартного отклонения квантиля выборки в каждом варианте. При увеличении одного варианта, предположим, что в этом варианте есть:

Участники  $i = 1, 2, 3, \dots, n$

Страницы просмотренные  $i$ -ым участником  $j = 1, 2, 3, \dots, P_i$ , где  $P_i$  (i.i.d) случайные значения удовлетворяющие функции распределения  $\mathcal{P}$ .

Загрузка  $j$ -ой страницы  $i$ -ым участником =  $X_{ij}$

$q$ -ый выборочный квантиль  $\{X_{ij}, i = 1, 2, \dots, n; j = 1, 2, 3, \dots, P_i\}$  обозначим как  $\hat{Q}$

Дисперсия -  $Var(\hat{Q})$ ; стандартное отклонение  $stddev(\hat{Q})$

First we define  $Y^{(n)}(x) = \frac{1}{n} \sum_i^n \sum_j^{P_i} \mathbb{I}_{\{X_{i,j} \leq x\}} = \frac{1}{n} \sum_i^n J_i$  and  $P^{(n)} = \frac{1}{n} \sum_{i=1}^n P_i$

, where  $J_i = \sum_{j=1}^{P_i} \mathbb{I}_{\{X_{i,j} \leq x\}}$ . Naturally  $J_i = 0$  if  $P_i = 0$ .

Under multidimensional central limit theorem,

$$\sqrt{n} \left( \begin{pmatrix} Y^{(n)}(x) \\ P^{(n)} \end{pmatrix} - \begin{pmatrix} \mu_J \\ \mu_P \end{pmatrix} \right) \xrightarrow{D} \mathcal{N}(0, \Sigma) \quad (1)$$

where  $\Sigma$  is the variance-covariance matrix of  $(Y^{(n)}(x), P^{(n)})$ ,  $\mu_J = \mathbb{E} \left( \sum_{j=1}^{P_i} \mathbb{I}_{\{X_{i,j} \leq x\}} \right) = \mathbb{E} \left[ \mathbb{E} \left( \sum_{j=1}^{P_i} \mathbb{I}_{\{X_{i,j} \leq x\}} | P_i \right) \right] = \mu_P F(x)$ ,  $\mu_P = \mathbb{E}[P_i]$  and  $F(x)$  is the cumulative distribution function of distribution  $\mathcal{F}$ .

Using the Delta method (Oehlert, 1992),

$$\sqrt{n} \left( \frac{Y^{(n)}(x)}{P^{(n)}} - \frac{\mu_J}{\mu_P} \right) \xrightarrow{D} \mathcal{N}(0, \sigma_{P,J}^2) \quad (2)$$

where  $\sigma_{P,J}^2 = \left( \frac{\mu_J}{\mu_P} \right)^2 \left( \frac{\Sigma_{JJ}}{(\mu_J)^2} + \frac{\Sigma^{PP}}{(\mu_P)^2} - 2 \frac{\Sigma^{PJ}}{\mu_J \mu_P} \right)$  with  $\Sigma^{JJ}$ ,  $\Sigma^{PP}$  and  $\Sigma^{PJ}$  elements in the  $2 \times 2$  variance-covariance matrix  $\Sigma$ .

Let  $F_n(x) = \frac{Y^{(n)}(x)}{P^{(n)}}$ , then the above expression can be written as,

$$\sqrt{n} (F_n(x) - F(x)) \xrightarrow{D} \mathcal{N}(0, \sigma_{P,J}^2) \quad (3)$$

When  $x = \hat{Q}$  the  $q$ -th sample quantile, that is,  $q = F_n(\hat{Q})$ ,

$$\sqrt{n} (q - F(\hat{Q})) \xrightarrow{D} \mathcal{N}(0, \sigma_{P,J}^2) \quad (4)$$

Applying the Delta method again, because  $\hat{Q}$  is a consistent estimate of  $Q$  the population quantile,

$$\sqrt{n} (F^{-1}(q) - \hat{Q}) \xrightarrow{D} \mathcal{N} \left( 0, \frac{\sigma_{P,J}^2}{f_X(Q)^2} \right) \quad (5)$$

Because  $F^{-1}(q) = Q$  and the standardized normal distribution is symmetric,

$$\sqrt{n} (\hat{Q} - Q) \xrightarrow{D} \mathcal{N} \left( 0, \frac{\sigma_{P,J}^2}{f_X(Q)^2} \right) \quad (6)$$

--2

// тут должна быть аннотация вставки

Таким образом асимптотическая оценка дисперсии квантиля это  $\frac{\sigma_{P,J}^2}{nf_X(Q)^2}$ , где  $Q$  можно оценить со средней плотностью в небольшом доверительном интервале  $\hat{Q}$ .

Предположим, что из  $n$  участников только  $i = 1, 2, \dots, n_0$  имеет ненулевые просмотры на интересующей странице.

Определим:

$$\mu_0^J = E(J_i | i = 1, 2, \dots, n_0); \mu_0^P = E(P_i | i = 1, 2, \dots, n_0); \Sigma_0 = Cov(J_i, P_i | i = 1, 2, \dots, n_0)$$



Тогда:

$$\begin{aligned}\mu^J &= \frac{n_0}{n} \mu_0^J, \mu^P = \frac{n_0}{n} \mu_0^P \\ \Sigma^{JJ} &= \frac{n_0}{n} \Sigma_0^{JJ} + \frac{n_0}{n} \left(1 - \frac{n_0}{n}\right) (\mu_0^J)^2 \\ \Sigma^{PP} &= \frac{n_0}{n} \Sigma_0^{PP} + \frac{n_0}{n} \left(1 - \frac{n_0}{n}\right) (\mu_0^P)^2\end{aligned}$$

$$\begin{aligned}\Sigma^{JP} &= \frac{n_0}{n} \Sigma_0^{JP} + \frac{n_0}{n} \left(1 - \frac{n_0}{n}\right) \mu_0^J \mu_0^P \\ \frac{1}{n} \sigma_{P,J}^2 &= \frac{1}{n} \left(\frac{\mu^J}{\mu^P}\right)^2 \left(\frac{\Sigma^{JJ}}{(\mu^J)^2} + \frac{\Sigma^{PP}}{(\mu^P)^2} - 2 \frac{\Sigma^{PJ}}{\mu^J \mu^P}\right) \\ &= \frac{1}{n} \left(\frac{\mu_0^J}{\mu_0^P}\right)^2 \left(\frac{\Sigma_0^{JJ}}{\frac{n_0}{n} (\mu_0^J)^2} + \frac{1 - \frac{n_0}{n}}{\frac{n_0}{n}} + \frac{\Sigma_0^{PP}}{\frac{n_0}{n} (\mu_0^P)^2} + \frac{1 - \frac{n_0}{n}}{\frac{n_0}{n}} - 2 \frac{\Sigma_0^{PJ}}{\frac{n_0}{n} \mu_0^J \mu_0^P} - 2 \frac{1 - \frac{n_0}{n}}{\frac{n_0}{n}}\right) \\ &= \frac{1}{n_0} \left(\frac{\mu_0^J}{\mu_0^P}\right)^2 \left(\frac{\Sigma_0^{JJ}}{(\mu_0^J)^2} + \frac{\Sigma_0^{PP}}{(\mu_0^P)^2} - 2 \frac{\Sigma_0^{PJ}}{\mu_0^J \mu_0^P}\right)\end{aligned}$$