**Блочный рандомизированный эксперимент**

[Ранее](http://baguzin.ru/wp/?p=5832) был рассмотрен *F*-критерий однофакторного дисперсионного анализа для оценки разностей между математическими ожиданиями *с* групп. Этот критерий применяется в ситуациях, когда *n* однородных элементов (так называемых экспериментальных объектов) случайным образом распределяются по *с* уровням исследуемого фактора (так называемые группы условий факторного эксперимента. Такие эксперименты называются полностью рандомизированными.[[1]](#footnote-1)

Кроме того, в разделе 9.2 описан ^-критерий для оценки разностей между математическими ожиданиями, который используется в ситуациях, связанных с повторяющимися измерениями или согласованными выборками. Этот критерий позволяет оценить различия между условиями проведения двух экспериментов. Предположим, что мы исследуем несколько групп условий или уровней исследуемого фактора. В таких ситуациях совокупности неоднородных объектов или индивидуумов, подлежащих сравнению (или повторным измерениям), называются блоками (blocks). Допустим, мы получили числовые результаты измерений для каждой группы условий и комбинаций блоков.

Эксперименты, в которых используются блоки, называются блочными рандомизированными экспериментами (randomized block designs). Хотя в таких схемах используются как условия, так и блоки, основное внимание уделяется оценке разностей между с разными группами условий. Целью объединения условий в блоки является максимально возможное исключение изменчивости экспериментальной ошибки с тем, чтобы разности между с групп условий проявились как можно отчетливее. Блочные рандомизированные эксперименты часто оказываются более эффективными, чем полностью рандомизированные эксперименты и, следовательно, позволяют получать более точные результаты [1,4, 7 и 8].

Для сравнения полностью рандомизированных и блочных рандомизированных экспе-риментов вернемся к сценарию, посвященному компании Perfect Parachute Company. Предположим, что в полностью рандомизированном эксперименте используются 12 наблюдений — по одному виду ткани на каждую из 12 смен. Любая изменчивость результатов испытаний становится частью экспериментальной ошибки, и, следовательно, различия между четырьмя поставщиками труднее уловить. Для того чтобы уменьшить экспериментальную ошибку, разработаем блочный рандомизированный эксперимент, в котором исследуются три смены, в течение каждой из которых ткутся четыре парашюта (один парашют — из волокон, полученных от первого поставщика, второй — из волокон, полученных от второго поставщика, и т.д.). Три смены рассматриваются как блоки, а условием факторного эксперимента является поставщик. Преимущество блочного рандомизированного эксперимента заключается в том, что из экспериментальной ошибки исключается изменчивость между тремя сменами. Следовательно, этот эксперимент часто обеспечивает более точные оценки различий между четырьмя поставщиками.

Критерии для оценки эффектов условий факторного эксперимента и блоков

Напомним, что в полностью рандомизированном эксперименте полная вариация (SST) подразделяется на межгрупповую (SSA) и внутригрупповую (SSW). Внутригрупповая вариация считается экспериментальной ошибкой, а межгрупповая вариация возникает вследствие различий между условиями факторного эксперимента.

Для того чтобы отделить эффект блокировки от экспериментальной ошибки блочно-

го рандомизированного эксперимента, необходимо подразделить внутригрупповую ва-

риацию на межблочную вариацию (SSBL) и случайную ошибку (SSE). Следова-

тельно, как показано на рис. 10.18, в блочном рандомизированном эксперименте

полная вариация результатов измерений представляет собой сумму межгрупповой

вариации (SSA), межблочной вариации (SSBL) и случайной ошибки (SSE).

Разделение полной вариации SST = SSA + SSBL + SSE

Случайная вариация (SSE) d.f. = (г- 1)(с- 1)

Рис. 10.18. Разделение полной вариации в блочном рандомизированном эксперименте

Для того чтобы разработать процедуру дисперсионного анализа для блочного рандо-мизированного эксперимента, введем следующие обозначения:

г — количество блоков,

с — количество групп или уровней фактора, п — общее количество наблюдений (п = гс), Хц — величина в i-м блоке и /-группе, Х1ш — среднее всех величин из i-ro блока,

Х^ — среднее всех величин из у-й группы,

с г

^^ГХу —общая сумма.

7=1 1=1

Полная вариация, называемая также полной суммой квадратов (SST), представляет собой вариацию между всеми наблюдениями. Величина SST равна сумме квадратов

разностей между каждым отдельным наблюдением и общим средним X , вычисленным по всем п наблюдениям.

ПОЛНАЯ ВАРИАЦИЯ

7=1 ,=1 (10.18)

где А" = —УУ J( — общее среднее значение.

Межгрупповая вариация, называемая также межгрупповой суммой квадратов (SSA), равна сумме квадратов разностей между выборочным средним каждой группы X t и общим средним значением X , деленным на количество блоков г.

МЕЖГРУППОВАЯ ВАРИАЦИЯ

SSA = r^X t-xJ, (10.19)

где X ( = -^Хп — среднее значение по у'-й группе.

Межблоковая вариация, называемая также межгрупповой суммой квадратов (SSBL), равна сумме квадратов разностей между средними значениями по каждому

блоку X, и общим средним значением X , деленному на количество групп с.

 МЕЖБЛОКОВАЯ ВАРИАЦИЯ

 SSBL = c^Xt -!)", (10.20)

где X, — среднее значение по i-му блоку.

Чисто случайная вариация или ошибка, также называемая суммой квадратов ошибок (SSE), равна сумме квадратов разностей между всеми наблюдениями после определенного воздействия и средними по блокам и группам.

СЛУЧАЙНАЯ ОШИБКА

sse=xzK - - + Щ- <10-21)

Поскольку фактор имеет с уровней, существует с-1 степеней свободы, связанных с межгрупповой суммой квадратов (SSA). Аналогично, поскольку существует г блоков, существует г-1 степеней свободы, связанных с межблоковой суммой квадратов (SSBL). Более того, общая сумма квадратов (SST) имеет п-1 степеней свободы, поскольку каждое наблюдение Хч сравнивается с общим средним X , вычисленным по всем п наблюдениям. Поскольку количество степеней свободы каждого источника вариации складывается с количеством степеней свободы полной вариации, количество степеней свободы суммы квадратов ошибок (SSE) получается путем вычитания и алгебраических манипуляций. Это количество степеней свободы равно (г-1)(с-1).

Если каждую компоненту суммы квадратов поделить на соответствующее количество степеней свободы, мы получим три вида дисперсии (MSA, MSBL и MSE), необходимых для формулы (10.22, а-в).

ВЫЧИСЛЕНИЕ СРЕДНЕКВАДРАТИЧЕСКИХ ЗНАЧЕНИИ

сел

№А = — , (10.22, а)

с-1

MSBL = ^^, (10.22,6)

г-1

MSE = 7—lb ГГ- (10.22, в)

Если выполняются предположения, принятые в дисперсионном анализе, можно применить .F-критерий (10.23), позволяющий проверить нулевую и альтернативную гипотезы о разностях между математическими ожиданиями с генеральных совокупностей.

if0: ц, = ц2 = ... = uf (условия не имеют эффекта), if,: не все ц, равны между собой, j = 1, 2, с.

РАНДОМИЗИРОВАННАЯ БЛОЧНАЯ F-СТАТИСТИКА ДЛЯ РАЗНОСТЕЙ МЕЖДУ С МАТЕМАТИЧЕСКИМИ ОЖИДАНИЯМИ

F=\*™ (10.23)

MSE

^-статистика имеет ^-распределение, в котором числитель MSA имеет с-1 степеней свободы, а знаменатель MSE имеет (г-1)(с-1) степеней свободы. При заданном уровне значимости а нулевая гипотеза отклоняется, если вычисленная ^-статистика больше верхнего критического значения Fu, присущего ^-распределению с с-1 и (г-1)(с-1) степенями свободы в числителе и знаменателе (см. табл. Д.5). Итак, решающее правило принимает следующий вид.

Нулевая гипотеза if0 отклоняется, если F > Fv; в противном случае гипотеза Н0 не отклоняется.

Для того чтобы выяснить, дает ли какие-либо преимущества блочный рандомизированный эксперимент, некоторые статистики предлагают применять F-критерий для проверки блоковых эффектов. Нулевая гипотеза заключается в отсутствии блоковых эффектов

ff0: цх = u2 = ... = и,, (блоковые эффекты не наблюдаются), if,: не все ц, равны между собой,j = 1, 2, г.

F-СТАТИСТИКА ДЛЯ БЛОКОВЫХ ЭФФЕКТОВ

р \_ MSBL ~ MSE

(10.24)

При заданном уровне значимости а нулевая гипотеза отклоняется, если вычисленная ^-статистика больше верхнего критического значения Fv, присущего F-pacnpeделению с с-1 и (г-1)(с-1) степенями свободы в числителе и знаменателе (см. табл. Д.5). Итак, решающее правило принимает следующий вид.

Нулевая гипотеза Н0 отклоняется, если F > Fv; в противном случае гипотеза ЬГ0 не отклоняется.

Некоторые статистики полагают, что этот критерий излишен, поскольку единственной целью блоков является создание более эффективного способа проверки наличия эффектов путем уменьшения экспериментальной ошибки.

В разделе 10.1 результаты дисперсионного анализа представлены в виде сводной таблицы ANOVA.

Таблица 10.8. Сводная таблица дисперсионного анализа для блочного рандомизированного эксперимента

Вид величины Количество Суммы Дисперсии F-статистика

степеней свободы квадратов

Межгрупповая с-1 SSA MSA=SSA/(c-l) F-MSA/MSW

Межблоковая г-1 SSBL MSBL=SSBL/(r-l) F=MSBL / MSE

Ошибка (r-2)(c-l) SSE MSE=SSE/(r-l)(c-l)

Полная rc-1 SST

Проиллюстрируем блочный рандомизированный эксперимент следующим примером. Предположим, что сеть ресторанов быстрого питания, имеющая четыре подразделения в определенном географическом регионе, желает оценить качество обслуживания в этих ресторанах. Для этой цели директор нанял шесть экспертов, имеющих разный опыт. Чтобы уменьшить эффект вариации между экспертами, был разработан блочный рандомизированный эксперимент, в которых блоками считались эксперты. В свою очередь, четыре ресторана образовали группы условий факторного эксперимента.

Шесть экспертов в случайном порядке инспектировали каждый из четырех ресторанов. Для оценки использовалась шкала баллов от 0 (низшая оценка) до 100 (высшая оценка). Результаты приведены в табл. 10.9.

Таблица 10.9. Рейтинги четырех ресторанов быстрого питания

Рестораны

Эксперты А Б В Г Всего Средние

1 70 61 82 74 287 71,75

2 77 75 88 76 316 79,00

3 76 67 90 80 313 78,25

4 80 63 96 76 315 78,75

5 84 66 92 84 326 81,50

6 78 68 98 86 330 82,50

Всего 465 400 546 476 1 887

Средние 77,50 66,67 91,00 79,33 78,625

Кроме того, как следует из табл. 10.9,

г = 6, с = 4, п = гс = 24

и

f = 1УУ У =1Ж = 78>625. rcjlTt 24

Результаты анализа результатов, полученных в рамках блочного рандомизированного эксперимента, приведены на рис. 10.19.

—

А .1.. а.

ИТОГИ

Счет Сумма Среднее Дисперсия

I Двухфакторный дисперсионный анализ без повторений

311

4 iЭксперт 1

5 ^Эксперт 2

6 'Эксперт 3

7 :Эксперт4

8 j Эксперт 5

9 :Эксперт 6 10]

II ;Ресторан А

12 Ресторан В

13 Ресторан С 14IРесторан D

сионныи анализ

15': »6!

17 ;Диспер

18 I

19]Строки

20 ■ Столбцы

21 i Погрешность

22'

23 jИтого

Источник вариации

SS

283,375 1787,458333 224,7916667

2295,625

MS

Р-значение F крит

5 56,675 3,781835032 0,020455782 2,901295204

3 595,8194444 39,75810936 2.23345Е-07 3,28738281 15 14,98611111

Рис. 10.19. Результаты инспекции сети ресторанов быстрого питания, полученные с помощью программы Microsoft Excel

Если установить уровень значимости критерия для проверки гипотезы о существовании различий между подразделениями сети ресторанов равным 0,05, решающее правило примет следующий вид: нулевая гипотеза Н0 (р, = р2 = р3 = р4) отклоняется, если F > 3,29. Число 3,29 представляет собой верхнее критическое значение F-распределения, имеющего три степени свободы в числителе и 15 степеней свободы в знаменателе (см. рис. 10.20). Поскольку F = 39,758 >FV = 3,29 (т.е. р = 0,000 < 0,05), мы можем отклонить гипотезу Н0 и утверждать, что средние рейтинги ресторанов статистически значимо различаются между собой. Чрезвычайно малое значение р означает, что, если бы средние рейтинги четырех подразделений были одинаковыми, вероятность обнаружить разности между их выборочными средними, была бы крайне малой. Итак, нулевая гипотеза практически невероятна. Следовательно, альтернативную гипотезу можно считать корректной.

Рис. 10.20. Области отклонения и принятия гипотез при изучении сети ресторанов быстрого питания при уровне значимости, равном 0,05, стремя и 15 степенями свободы

Рис. 10.21. Области отклонения и принятия гипотез при изучении сети ресторанов быстрого питания при уровне значимости, равном 0,05, с пятью и 15 степенями свободы

Для проверки эффективности блокировки, можно проверить разность между экспертами. При 5% -м уровне значимости решающее правило можно сформулировать следующим образом: нулевая гипотеза if0(p1 = р2 = ... = р6) отклоняется, если вычисленная

статистика F>2,90. Число 2,90 представляет собой верхнее критическое значение ^-распределения, имеющего пять степеней свободы в числителе и 15 степеней свободы в знаменателе (см. рис. 10.21). Поскольку F = 3,782 > Fv = 2,90 (т.е. р = 0,02 < 0,05), мы может отклонить гипотезу Н0 и утверждать, что средние рейтинги ресторанов статистически значимо различаются между собой. Итак, применение блоков уменьшает экспериментальную ошибку.

Процедуры Excel: дисперсионный анализ с помощью рандомизированного эксперимента

Чтобы выполнить однофакторный дисперсионный анализ, следует применить процедуру I

Анализ данных...^Двухфакторный дисперсионный анализ без повторений. В надстройке PHStat2 j

эта процедура не предусмотрена. Кроме того, вследствие сложности вычислений шаблон рабочего I

листа для этого критерия довольно трудно реализовать вручную. 1

Например, чтобы осуществить дисперсионный анализ данных, приведенных в табл. 10.9, используя блочный рандомизированный эксперимент, необходимо открыть рабочий лист Рейтинги в рабочей книге Chapter 10.xls и выполнить такие действия.

Двухфакторный дисперсионный анализ без повторений |Х[

Входные данные Входной интервал:

0 Метки Альфа: :о,05 \

Параметры вывода О Выходной интервал: (\*} Новый рабочий лист: О Новая рабочая книга

Выбрать команду Сервисе Анализ данных...

\ Отмена j Справка

В диалоговом окне Анализ данных выбрать пункт Двухфакторный дисперсионный анализ без повторений в списке Инструменты анализа. Щелкнуть на кнопке ОК.

В диалоговом окне Двухфакторный дисперсионный анализ без повторений

(см. иллюстрацию) сделать следующее.

3.1. Ввести в окне редактирования Входной интервал переменной 1

3.2. 3.3. 3.4.

3.5.

диапазон А1:Е7.

Установить флажок Метки.

Ввести в окне редактирования Альфа число 0, 05.

Установить переключатель Параметры вывода в положение Новый рабочий лист и ввести название нового листа.

Щелкнуть на кнопке ОК.

Рабочий лист, созданный с помощью этой процедуры, не является динамически обновляемым. Следовательно, если данные изменятся, все описанные выше действия необходимо повторить.

Для выполнения этой процедуры необходимо, чтобы данные для каждой группы располагались в разных столбцах. Такие данные называются разгруппированными. Для того чтобы обработать сгруппированные данные, следует выполнить процедуру, описанную в разделе ЕН.9.2.

Кроме обычных ограничений, принятых в однофакторном дисперсионном анализе, необходимо также предположить, что между условиями факторного эксперимента и блоками нет взаимодействия. Иначе говоря, необходимо, чтобы все различия между условиями эксперимента (ресторанами) были согласованы со всеми блоками (отмечались всеми экспертами). Понятие взаимодействия (interaction) обсуждается в разделе 10.2.

После разработки схемы блочного рандомизированного эксперимента и анализа данных о рейтингах ресторанов возникает вопрос: какой эффект оказывает блокирование на дисперсионный анализ? Иначе говоря, получаем ли мы более точные результаты, применяя блокирование при анализе разных групп условий факторного эксперимента? Для того чтобы ответить на этот вопрос, следует вычислить оценку относите льной эффективности (relative efficiency — RE) блочного рандомизированного эксперимента по сравнению с полностью рандомизированным экспериментом.

Г ОЦЕНКА ОТНОСИТЕЛЬНОЙ ЭФФЕКТИВНОСТИ

(r-\)MSBL + r(c-\)MSE

I RE = ± '— —\* '- . (10.25)

j (rc-\)MSE

Используя формулу (10.19), получаем

5x56,675 + 6x3x14,986 , \_л

Kb = = 1,60.

23x14,986

Это означает, что для получения такой же точности при сравнении средних по группам в рамках однофакторного дисперсионного анализа нам понадобилось бы в 1,6 раза увеличить количество наблюдений в каждой группе.

Множественные сравнения: процедура Тьюки

Как и в полностью рандомизированном эксперименте, отклонив нулевую гипотезу о равенстве всех средних по группам, мы можем определить, какая группа условий значительно отличается от остальных. Для блочного рандомизированного эксперимента такая процедура была разработана Джоном Тьюки (John Tukey) [7-9]. Критический размах в процедуре Тьюки (Tukey procedure) вычисляется по формуле (10.26).

КРИТИЧЕСКИЙ РАЗМАХ

Критический размах — Qv^ ? (10.26)

где статистика Qv представляет собой верхнее критическое значение распределения стьюдентизованного размаха, имеющего с степеней свободы в числителе и (г-1)(с-1) степеней свободы в знаменателе. Величины распределения стьюдентизованного размаха приведены в табл. Д.9.

Каждая из с(с-1)/2 пар средних сравнивается с одним критическим размахом. Пара, например, группа у— группа у" объявляется статистически значимо разными, если модуль разности между выборочными средними JA^-X^I превышает критический размах.

Продемонстрируем применение процедуры Тьюки на примере анализа сети ресторанов. Поскольку проверке подвергаются четыре ресторана, в процедуре Тьюки будет выполнено 4(4-1)/2 = 6 попарных сравнений. Из данных, приведенных на рис. 10.19, следует, что модули разностей принимают перечисленные ниже значения.

1. 1Х| -х2\ = |77,50 -6б,67| = 10,83.

2. 1\*, -х\ = |77,50 -91,00) = 13,50.

3. г\*' -\*л\ = |77,50 -79,33| = 1,83.

4. \Х2 -х, = |б6,67 -91,00| = 24,33.

5. -хА = |б6,67 -79,33| = 12,66.

6. \х, -х4 = |91,00 -79,33| = 11,67.

Для того чтобы вычислить критический размах, определим по рис. 10.19 величины MSE = 14,986 и г= 6. По табл. Д.9 (для а = 0,05, с = 4и (г-1)(с-1) = 15) находим, что величина Q„ — верхнее критическое значение тестовой статистики, имеющей четыре степени свободы в числителе и 15 степеней свободы в знаменателе— равна 4,08. Используя формулу (10.26), получаем

Обратите внимание на то, что все попарные разности, за исключением величины |х, -X^ , превышают критический размах. Следовательно, между всеми ресторанами,

за исключением ресторанов А и Г, существуют значительные отличия. Кроме того, под-разделение В имеет наивысший рейтинг (т.е. работает лучше остальных), а подразделение Б — наименьший (т.е. работает хуже всех).

Illliill

111111

Изучение основ

10.26. Предположим, что в блочном рандомизированном эксперименте изучается один

фактор, имеющий пять уровней и семь блоков.

1. Сколько степеней свободы существует при определении межгрупповой вариации?

2. Сколько степеней свободы существует при определении межблоковой вариации?

3. Сколько степеней свободы существует при определении случайной вариации, или ошибки?

4. Сколько степеней свободы существует при определении полной вариации?

10.27. Вернемся к задаче 10.26.

1. Чему равна величина SSE, если SSA = 60, SSBL = 75 и SST = 210?

2. Чему равна величина MSA1

3. Чему равна величина MSBL?

4. Чему равна величина MSE?

5. Чему равна тестовая статистика F для оценки разностей между пятью средними?

6. Чему равна тестовая статистика F для оценки блочных эффектов?

10.28. Вернемся к задачам 10.26 и 10.27.

1. Сформируйте и заполните сводную таблицу дисперсионного анализа.

2. Чему равно верхнее критическое значение F-распределения при оценке раз-

ностей между пятью средними с уровнем значимости, равным 0,05?

3. Сформулируйте решающее правило для проверки нулевой гипотезы о том, что все пять групп имеют одинаковые средние значения.

4. Какое статистическое решение вы примете?

5. Чему равно верхнее критическое значение ^-распределения при оценке блочных эффектов с уровнем значимости, равным 0,05?

6. Сформулируйте решающее правило для проверки нулевой гипотезы о том, что блочных эффектов нет.

7. Какое статистическое решение вы примете?

10.29. Вернемся к задачам 10.26-10.28.

1. Сколько степеней свободы в числителе и знаменателе имеет распределение стьюдентизованного размаха при выполнении процедуры Тьюки?

2. Чему равно верхнее критическое значение распределение стьюдентизованного размаха при уровне значимости, равном 0,05?

3. Объясните смысл критического размаха в процедуре Тьюки.

10.30. Предположим, что в блочном рандомизированном эксперименте изучаются один

фактор, три уровня и семь блоков.

1. Сколько степеней свободы существует при определении межгрупповой вариации?

2. Сколько степеней свободы существует при определении межблоковой вариации?

3. Сколько степеней свободы существует при определении случайной вариации, или ошибки?

4. Сколько степеней свободы существует при определении полной вариации?

10.31. Вернемся к задаче 10.30. Предположим, что величина SSA = 36, а рандомизиро-

ванная блочная F-статистика равна 6,0.

1. Чему равна величина MSE?

2. Чему равна величина SSE?

3. Чему равна величина SSBL, если F-статистика для оценки блочных эффектов равна 4,0?

4. Чему равна величина SST?

5. Существует ли эффект условий, если уровень значимости равен 0,01?

6. Существует ли эффект блоков, если уровень значимости равен 0,01?

10.32. Ниже приведена неполная сводная таблица дисперсионного анализа блочного

рандомизированного эксперимента, имеющего четыре уровня и восемь блоков.

Заполните недостающие ячейки.

Вид величины Количество Суммы Дисперсии F-статистика

 степеней свободы квадратов

Межгрупповая с-1 = ? SSA- ? MSA = 80 F = ?

Межблоковая г-1 = ? SSBL = 540 MSBL = ? F- 5,0

Ошибка (г-1)(с-1) = ? SSE = ? MSE = ?

Полная rc-1 = ? SST- ?

10.33. Вернемся к задаче 10.32.

1. Чему равно верхнее критическое значение F-распределения при оценке разностей между четырьмя средними с уровнем значимости, равным 0,05?

2. Сформулируйте решающее правило для проверки нулевой гипотезы о том, что все четыре группы имеют одинаковые математические ожидания.

3. Какое статистическое решение вы примете?

4. Чему равно верхнее критическое значение F-распределения при оценке блочных эффектов с уровнем значимости, равным 0,05?

5. Сформулируйте решающее правило для проверки нулевой гипотезы о том, что блочных эффектов нет.

6. Какое статистическое решение вы примете?

Применение понятий

Примечание: рекомендуем для решения задач 10.34-10.38 использовать программу Microsoft Excel.

10.34. Девять экспертов были приглашены для дегустации четырех сортов колумбийского кофе. Для того чтобы обеспечить объективность оценки, каждый из девяти дегустаторов испытывал каждый сорт кофе в случайном порядке. Четыре характеристики кофе — вкус, аромат, насыщенность и кислотность — оценивались по 7-балльной шкале (1 — очень плохо, 7 — очень хорошо). Полученные результаты приведены в следующей таблице. Ђ#C0FFEE . XLS.

Марка

Эксперт А Б В Г

С.С. 24 26 25 22

S.E. 27 27 26 24

E.G. 19 22 20 16

B.L. 24 27 25 23

СМ. 22 25 22 21

C.N. 26 27 24 24

G.N. 27 26 22 23

R.M. 25 27 24 21

P.V. 22 23 20 19

Проанализируйте данные и определите, существуют ли различия между суммарными рейтингами четырех сортов колумбийского кофе при уровне значимости, равном 0,05. Если различия есть, определите, какой из сортов кофе получил наивысшую оценку. Как вы пришли к такому выводу? 10.35. Группа студентов, изучающих бизнес-статистику, поставила эксперимент, в ходе которого исследовался вопрос: влияет ли марка жевательной резинки на размер шарика, который можно из нее выдуть. Студенты были уверены, что Кайл является экспертом по жевательным резинкам, и поэтому его опыт мог отрицательно повлиять на результаты полностью рандомизированного эксперимента. Итак, чтобы уменьшить межличностную изменчивость, студенты решили применить схему блочного рандомизированного эксперимента, в котором блоками являлись бы они сами. Студент разжевывал два кусочка жевательной резинки определенной марки, а затем выдувал два шарика, пытаясь раздуть их как можно больше. Другой студент измерял максимальные диаметры шариков. В следующей таблице приведены результаты 16 наблюдений (в дюймах). ^BUBBLEGUM. XLS.

Марка жевательной резинки

Студент Bazooka Bubbletape Bubbleyum Bubblicious

Кайл 8,75 9,50 8.50 11,50

Сара 9,50 4,00 8,50 11,00

Лей 9,25 5,50 7,50 7,50

Исаак 9,50 8,50 7,50 7,50

1. Можно ли утверждать, что средние диаметры шариков, полученных из жева-тельных резинок, различаются между собой?

2. Если возможно, примените процедуру Тьюки и определите, какая из марок жевательных резинок отличается от остальных. (Уровень значимости а установите равным 0,05.)

3. Считаете ли вы, что в данном эксперименте наблюдается значительный блочный эффект? Обоснуйте свой ответ.

4. Считаете ли вы, что Кайл действительно лучше остальных умеет выдувать шарики из жевательных резинок?

10.36. Менеджер крупного агентства по торговле недвижимостью закончил обучение

трех вновь нанятых работников методам оценки недвижимости. Для того чтобы

оценить эффективность своего метода обучения, менеджер решил определить,

существует ли какая-либо разница между оценками недвижимости, выставлен-

ными каждым из новых агентов. Менеджер случайным образом выбрал 12 домов

и поручил каждому агенту оценить эти дома (в тыс. долл.) Результаты приведе-

ны в файлеIIREAPPR3.XLS.

1. Примените блочный рандомизированный эксперимент и определите, существует ли разница между средними оценками, выставленными тремя агентами.

2. Какие условия должны выполняться, чтобы можно было осуществить такую проверку?

3. К каким выводам пришел менеджер? Эффективен ли его метод обучения агентов. Одинаково ли они оценивают недвижимость? Обоснуйте свой ответ.

10.37. Компания Philips Semiconductors является ведущим европейским производителем

интегральных микросхем. Основой для микросхем являются кремниевые под-

ложки, которые предварительно доводят до требуемой толщины. Подложки уста-

навливают в разных местах шлифовального круга и фиксируют с помощью ваку-

умной декомпрессии. Одной из целей производственного процесса является

уменьшение изменчивости толщины подложки в зависимости от ее места на шли-

фовальном круге и партии. Были собраны данные о 30 партиях. В каждой партии

измерялась толщина подложек, находящихся на позициях 1 и 2 (внешняя часть

шлифовального круга), 18 и 19 (середина круга) и 28 (внутренняя часть круга). Ре-

зультаты приведены в файле CIRCUIT . XLS. Выполните полный анализ данных

при уровне значимости, равном 0,01, и определите, существуют ли различия меж-

ду средней толщиной подложек, расположенных на пяти позициях. Если такие

различия существуют, укажите, какие позиции отличаются друг от друга. Какой

вывод следует сделать?

Источник: К. С. В. Roes and R. J. M. M. Does, "Shewhart-type Charts in Nonstandard Situations," Technometrics, 37,1995,15-24.

10.38. Данные, приведенные в файле ^icONCRETE2 . XLS, описывают прочность на сжа-

тие (psi — тыс. фунтов на кв. дюйм) 40 образцов бетона, взятых на 2-, 7- и 28-й

дни после укладки.

Источник: О. Carillo-Gamboa and R.F.Gunst, "Measurement-Error-Model Collinearities", Technometrics, 34,1992, 454-464.

1. Можно ли утверждать, что между средней прочностью на сжатие образцов бетона, взятых на 2-, 7- и 28-й дни после укладки, существует значительное различие, если уровень значимости равен 0,05?

2. Если возможно, примените процедуру Тьюки и определите, какой день после укладки значительно отличается от остальных с точки зрения прочности бетона на сжатие. (Уровень значимости а установите равным 0,05.)

3. Определите относительную эффективность блочного рандомизированного эксперимента по сравнению с однофакторным полностью рандомизированным экспериментом.

РЕЗЮМЕ

В главе описана концепция полностью рандомизированного эксперимента. Приведены различные процедуры анализа эффекта одного и двух факторов. Рассмотрен практический пример, иллюстрирующий применение критериев одно- и двухфакторного дисперсионного анализа. Детально описаны условия выполнения статистических процедур. Напомним, что проверка необходимых условий является неотъемлемой частью статистического анализа и позволяет правильно выбирать критерии. Как показано на структурной схеме, существует несколько подходов к сравнению групп числовых данных на основе экспериментов.

F Планы | экспериментов

Полностью рандомизированный план

Факторный план

I

Критерий Левенэ

7 F-критерий для I однофакторного дисперсионного анализа

F-критерий для однофакторного дисперсионного анализа:

Множественное Г сравнение

Множественное сравнение

F-критерий

в однофакторном анализе, 646

для фактора А, 667

для фактора В, 667

для эффекта взаимодействия факторов А и В, 667 Вариация

взаимодействия, 666

внутригрупповая, 643; 645

межгрупповая, 643; 645

полная, 644; 665

фактора А, 665

фактора В, 665 Дисперсионный анализ, 643

двухфакторный, 664

однофакторный, 643 Дисперсия

MSA, 666

MSAB, 666

MSB, 666

MSE, 666 Критерий Левенэ, 656 Критический размах, 653 Множественное сравнение, 675 Общее среднее, 644 Ошибка

случайная, 666

эксперимента, 643 План

полностью рандомизированный, 642 факторный, 642 Процедура ANOVA

двухфакторная, 664

однофакторная, 643 апостериорного сравнения, 653 множественного сравнения, 653 Тьюки-Крамера, 653

Реплика, 664

Сводная таблица ANOVA, 647 Сумма квадратов

внутригрупповая, 645

межгрупповая, 645

ошибок, 666

полная, 645; 665

соответствующая фактору А, 665 соответствующая фактору В, 665 средняя, 646

учитывающая взаимодействие между факторами А и В, 665 Фактор, 642 Эксперимент

полностью рандомизированный, 642

факторный, 665 Эффект

взаимодействия, 670

главный, 670

условий эксперимента, 643

УПРАЖНШИЯ К ГЛАВЕ 10 \*\* :«' \*' !' Л '/^Л"< <V Л

Проверка знаний

10.39. Чем межгрупповая дисперсия MSA отличается от внутригрупповой дисперсии MSW?

10.40. В чем заключается разница между полностью рандомизированным и двухфак-торным экспериментами?

10.41. Какие условия необходимы для проведения дисперсионного анализа?

10.42. При каких условиях можно применять ^-критерий в однофакторном дисперсионном анализе для оценки разностей между математическими ожиданиями с генеральных совокупностей?

10.43. Когда и как следует применять процедуры множественного сравнения для попарного сопоставления математических ожиданий с генеральных совокупностей?

10.44. В чем заключается разница между однофакторным дисперсионным анализом и критерием Левенэ?

10.45. При каких условиях можно применять F-критерий в двухфакторном дисперсионном анализе для оценки разностей между математическими ожиданиями каждого фактора в факторном эксперименте?

10.46. В чем заключается взаимодействие между факторами в двухфакторном эксперименте?

10.47. Как применить F-критерий в двухфакторном дисперсионном анализе для оценки эффекта взаимодействия между факторами?

Применение понятий

Задачи 10.48-10.53 можно решать как вручную, так и с помощью программы Microsoft Excel.

10.48. Управляющий заводом, производящим бытовую технику, хотел бы определить

оптимальную продолжительность работы стиральной машины. Для этого он спла-

нировал эксперимент, в ходе которого учитывалось влияние марки стирального

порошка и продолжительности работы стиральной машины на качество стирки.

В эксперименте использовались четыре марки стирального порошка (А, В, С и D)

и четыре разных цикла стирки (18, 20, 22 и 24 мин.). В 32 стиральные машины за-

гружался одинаковый объем случайно выбранного одинаково загрязненного бе-

лья: по две загрузки на каждую из 16 комбинаций факторов. Результаты (вес уда-

ленной грязи в фунтах) представлены в таблице. ^LAUNDRY. XLS.

Продолжительность стирки (мин.)

Марка стирального порошка 18 20 22 24

А 0,11 0,13 0,17 0,17

 0,09 0,13 0,19 0,18

В 0,12 0,14 0,17 0,19

 0,10 0,15 0,18 0,17

С 0,08 0,16 0,18 0,20

 0,09 0,13 0,17 0,16

D 0,11 0,12 0,16 0,15

 0,13 0,15 0,17 0,17

1. Существует ли статистически значимый эффект взаимодействия между маркой стирального порошка и продолжительностью стирки, если уровень значимости равен 0,05?

2. Существует ли статистически значимый эффект стирального порошка, если уровень значимости равен 0,05?

3. Существует ли статистически значимый эффект продолжительности стирки, если уровень значимости равен 0,05?

4. Постройте график зависимости среднего объема удаленной грязи (в фунтах) от марки стирального порошка при фиксированной продолжительности стирки.

5. Если возможно, примените процедуру Тьюки и определите различия между марками стирального порошка и продолжительностью стирки.

6. Какую продолжительность стирки можно рекомендовать для данной стиральной машины?

7. Повторите анализ, считая продолжительность стирки единственным фактором. Сравните результаты с решениями задач 1-6.

10.49. Начальник ОТК на ткацкой фабрике хотел бы сравнить влияние мастерства ткачих и марки станка на прочность шерстяной ткани. Для этого ткань была разрезана на квадратные куски со стороной один ярд, которые были случайным образом распре-делены между группами: по три отреза на каждую из 12 комбинаций (четыре ткачихи и три станка). Результаты приведены в таблице. ^BREAKSTW. XLS.

Станок

Ткачиха 1 2 3

А 115 111 109

 115 108 110

 119 114 107

В 117 105 110

 114 102 113

 114 106 114

С 109 100 103

 110 103 102

 106 101 105

D 112 105 108

 115 107 111

 111 107 110

1. Существует ли статистически значимый эффект взаимодействия между ткачихами и марками станка, если уровень значимости равен 0,05?

2. Существует ли статистически значимый эффект ткачихи, если уровень значимости равен 0,05?

3. Существует ли статистически значимый эффект марки станка, если уровень значимости равен 0,05?

4. Постройте график зависимости средней прочности (в фунтах) от вида станка для каждой ткачихи.

5. Если возможно, примените процедуру Тьюки и определите различия между ткачихами и марками станка.

6. Влияет ли мастерство ткачихи и марка станка на прочность шерстяной ткани? Обоснуйте свой ответ.

7. Повторите анализ, оставив станки в качестве единственного фактора. Сравните результаты с решениями задач 3, 5 и 6.

10.50. Рассмотрим эксперимент, в ходе которого измеряется прочность пряжи. Эксперимент состоит из двух этапов.

Этап 1. Руководитель производства хотел бы знать, влияет ли давление воздуха (в фунтах на квадратный дюйм) на прочность пряжи. Рассматриваются три уровня давления: 30, 40 и 50 фунтов на квадратный дюйм. Для анализа из одной и той же партии отобраны 18 однородных экземпляров пряжи, которые затем распределяются по уровням давления: по шесть штук на каждый из трех уровней. Результаты приведены в файле YARN.XLS.

1. Каким условиям должна удовлетворять дисперсия показателей прочности, соответствующих трем уровням давления?

2. Существует ли статистически значимая разница между средними прочностями пряжи при разном давлении воздуха, если уровень значимости равен 0,05?

3. Если возможно, примените процедуру Тьюки-Крамера и определите, какой уровень давления статистически значимо влияет на прочность пряжи, если уровень значимости равен 0,05.

4. Какой вывод должен сделать руководитель производства?

Этап 2. Предположим, что руководитель производства может измерить не только давление воздуха (в фунтах на квадратный дюйм), но и направление его потока. Следовательно, вместо полностью рандомизированного эксперимента на втором этапе осуществляется двухфакторный эксперимент. Первый фактор — направление потока — имеет два уровня: попутный и встречный. Второй фактор — давление — по-прежнему имеет три уровня: 30, 40 и 50 фунтов на квадратный дюйм. Для анализа из одной и той же партии отобраны 18 однородных экземпляров пряжи, которые затем распределяются по уровням давления: по три экземпляра на каждую комбинацию факторов. Результаты приведены в таблице. llYARN. XLS.

Давление воздуха (фунты на кв. м.)

Направление 30 40 50

потока

 25,5 24,8 23,2

Попутный 24,9 23,7 23,7

 26,1 24,4 22,7

Давление воздуха (фунты на кв. м.)

Направление 30 40 50

потока

 24,7 23,6 22,6

Встречный 24,2 23,3 22,8

 23,6 21,4 24,9

5. Существует ли статистически значимый эффект взаимодействия между на-правлением потока и давлением воздуха, если уровень значимости равен 0,05?

6. Существует ли статистически значимый эффект направления потока, если уровень значимости равен 0,05?

7. Существует ли статистически значимый эффект давления воздуха, если уровень значимости равен 0,05?

8. Постройте график зависимости средней прочности пряжи для двух направлений потока при разных давлениях.

9. Если возможно, примените процедуру Тьюки и определите различия между разными величинами давления.

10. Какой вывод можно сделать на основе результатов решения задач 5-8?

11. Сравните результаты решения задач 5-10 с результатами решения задач 1-4.

10.51. Современное программное обеспечение требует все более быстрого доступа к данным. Рассмотрим эксперимент, в ходе которого оценивается влияние размера файла на скорость доступа.

Этап1. Рассматриваются три размера файла: небольшой— 50 000 символов, средний — 75 000 символов и большой — 100 000 символов. В эксперименте оценивались восемь файлов каждого размера. Результаты (в миллисекундах) приведены в файле ^ACCESS . XLS.

1. Каким условиям должна удовлетворять дисперсия скорости доступа, соот-ветствующей трем размерам файла?

2. Существует ли статистически значимая разница между средними скоростями доступа при разных размерах файла, если уровень значимости равен 0,05?

3. Если возможно, примените процедуру Тьюки-Крамера и определите, какой размер файла статистически значимо влияет на скорость доступа, если уровень значимости равен 0,05.

4. Какой вывод можно сделать на основе этих данных?

Этап 2. Предположим, что в эксперименте учитывается не только размер файла, но и размер буфера ввода-вывода. Следовательно, вместо полностью рандомизированного эксперимента на втором этапе осуществляется двухфакторный эксперимент. Первый фактор — размер буфера — имеет два уровня: 20 и 40 Кбайт. Второй фактор — размер файла — по-прежнему имеет три уровня: небольшой, средний и большой. Для анализа скорости доступа выполняются четыре программы (реплики) для каждой комбинации факторов. Результаты приведены в файле ^ACCESS . XLS.

5. Существует ли статистически значимый эффект взаимодействия между раз-

мерами файла и буфера, если уровень значимости равен 0,05?

6. Существует ли статистически значимый эффект размера буфера, если уровень значимости равен 0,05?

7. Существует ли статистически значимый эффект размера файла, если уровень значимости равен 0,05?

8. Постройте график зависимости средней скорости доступа от размера буфера при разных размерах файла. Опишите взаимодействие этих факторов и дайте интерпретацию их главных эффектов.

9. Какой вывод можно сделать на основе решений задач 5-8?

10. Сравните результаты решения задач 5-8 с решениями задач 1-4. 10.52. Группа студентов, изучающих статистику, провела эксперимент, в ходе которого измерялось время загрузки компьютеров трех разных типов (MAC, iMAC и Dell).

Этап 1. Студенты случайным образом выбирали один компьютер из каждой группы. Они заходили на игровой Web-сайт компании Microsoft и загружали компьютерную игру в баскетбол. Регистрировалось время между щелчком на ссылке и окончанием загрузки. После каждой загрузки файл удалялся, а мусорная корзина очищалась. Порядок 30 загрузок определялся случайным образом. ^COMPUTERS . XLS. Выполните анализ данных, приведенных в таблице.

Компьютер

MAC iMAC Dell

156 160 236

166 165 238

148 184 257

160 192 242

139 197 282

151 172 253

158 189 270

167 179 256

142 200 267

219 193 259

Этап 2. Во втором эксперименте, выполненном группой студентов, в качестве второго фактора исследовался браузер (Netscape Communicator и Interner Explorer). В этом эксперименте использовались только два вида компьютеров: MAC и Dell. Было выполнено по восемь загрузок при каждой из четырех комбинаций факторов. ft\*C0MPUTERS2 . XLS. Выполните анализ данных из следующей таблицы.

Компьютер

Браузер MAC Dell

 142 284

 132 304

 125 273

Netscape Communicator 136 340

 127 326

 138 301

 147 291

 143 285

Компьютер

Браузер MAC Dell

 198 285

 210 292

 199 305

Internet Explorer 202 325

 196 297

 213 301

 207 285

 201 290

Отчеты

10.53. В файле ^BEER.XLS приведены данные об упаковках, содержащих шесть 12-

унциевых бутылок пива 69 сортов. В их число входят цена, количество калорий

в 12 жидких унциях, процентное содержание алкоголя в 12 жидких унциях, вид

пива (светлое, эль, импортное легкое, обычное, ледяное, легкое, безалкоголь-

ное), а также страны производства (США или другие).

Напишите отчет, содержащий полную оценку каждой переменной — цены, количества калорий и содержания алкоголя для каждого вида пива — светлого, эля, импортного легкого, обычного, ледяного, легкого, безалкогольного. После этого выполните аналогичное исследование каждой числовой переменной, учитывая страну производства. Включите в отчет все необходимые таблицы, диаграммы и количественные показатели, полученные в ходе исследования.

Для создания отчетов, предложенных в этом разделе, воспользуйтесь программой Microsoft Excel и другими средствами пакета Microsoft Office (в частности, программой для подготовки презентаций PowerPoint).

Источник: "Beers". Copyright © 1996 by Consumers Union of U.S. Inc., Yonkers, N.Y. Цитируется с разрешения журнала Consumer reports, June 1996.

Применение Интернет

10.54. Зайдите на сайт www.prenhall.com/levine. Выберите ссылку Chapter 10

и щелкните на ссылке Internet exercises.

';\ ГРУППОВОЙ ПРОЕКТ

ТР.10.1. Файл данных ^MUTUAL FUNDS . XLS содержит информацию о 13 переменных, характеризующих 259 взаимных фондов.

Фонд — название взаимного фонда.

Вид — вид акций, принадлежащих взаимному фонду: малые, средние и крупные компании.

Цель — цель фонда (быстрый или медленный рост капитала). Активы — в млн. долл. Комиссия — да или нет.

Издержки — издержки, понесенные взаимным фондом (в процентах от среднего объема чистых активов).

Доходность 2 001 — доходность за двенадцать месяцев 2001 г.

Трехлетняя доходность — среднегодовая доходность за период с 1999 по 2001 гг.

Пятилетняя доходность — среднегодовая доходность за период с 1997 по 2001 гг.

Оборачиваемость — уровень торговой активности фонда: очень низкий, низкий, средний, высокий, очень высокий.

Риск — уровень риска: очень низкий, низкий, средний, высокий, очень высокий. Лучший квартал — квартал с наивысшей доходностью за период с 1997 по 2001 гг.

Худший квартал — квартал с наименьшей доходностью за период с 1997 по 2001 гг.

Оцените доходность взаимных фондов, относящихся к разным категориям (малые, средние и крупные компании) в 2001 году, за три года (с 1999 по 2001 г.) и за пять лет (с 1997 по 2001 г.) при уровне значимости, равном 0,05. Выполните анализ данных при уровне значимости, равном 0,05. Изложите свои выводы в письменном виде. Включите в отчет все необходимые таблицы, диаграммы и количественные показатели, полученные в ходе исследования. Для создания отчета используйте программу Microsoft Excel и другие средства пакета Microsoft Office (в частности, программу для подготовки презентаций PowerPoint).

Этап 1

Группа, разрабатывающая новую подписную политику, решила проверить, как влияет прямой маркетинг по телефону на количество подписчиков. После нескольких рабочих совещаний, в которых приняли участие как инструкторы, так и непосредственные участники телефонных рекламных акций, было принято решение повысить продолжительность телефонных переговоров с потенциальными подписчиками, поскольку при более продолжительном разговоре вероятность оформления подписки повышается.

Группа решила оценить влияние продолжительности телефонного разговора с по-тенциальным подписчиком на успех подписной кампании. Оказалось, что телефонные разговоры с 19:00 до 21:00 оказались значительно более продолжительными, чем разговоры с 17:00 до 19:00.

После этого группа решила исследовать влияние разновидности презентации на про-должительность разговора. Для этого была создана случайная группа из 24 женщин, раз-деленная на три подгруппы по восемь женщин. В каждой подгруппе использовалась структурированная, полуструктурированная и совершенно не структурированная презентация. Все звонки выполнялись с 19:00 до 21:00. Звонившие обязаны были обращаться к потенциальному подписчику неформально. (Например: "Привет, это Мэри Джонс из газеты Spingville Herald. Могу я поговорить с Биллом Ричардсом?".) Абоненты знали, что исследователи внимательно следят за их работой, но не знали, какой именно разговор будет прослушан. Продолжительность телефонных разговоров измерялась в секундах, прошедших от момента, когда абонент ответил на вопрос, до момента, когда он положил трубку. Результаты приведены в табл. SH.10.1. Ђ#SH10-1 .XLS.

Таблица SH.10.1. Продолжительность телефонного разговора (в секундах) в зависимости от плана презентации

Проанализируйте данные, приведенные в табл. SH. 10.1. Изложите свои выводы и рекомендации в отчете. В приложении обоснуйте свой выбор статистического критерия для сравнения трех независимых групп.

@ НЕ ПРОДОЛЖАЙТЕ, ПОКА НЕ ЗАКОНЧИТЕ ПЕРВЫЙ ЭТАП Этап 2

После анализа данных, представленных в табл. SH.10.1, становится очевидным, что структурированный план презентации позволяет значительно увеличить продолжительность разговора по сравнению с полуструктурированной и неструктурированной презентациями. Группа вновь собралась на совещание и стала искать новые возможности повысить продолжительность телефонных переговоров с потенциальными подписчиками и, следовательно, увеличить количество подписчиков. Оказалось, что структурированный телефонный разговор, проведенный с 19:00 до 21:00, наиболее эффективен. В поисках новых возможностей группа решила оценить два дополнительных фактора, влияющих на продолжительность телефонной презентации.

• Пол абонента: мужской или женский.

• Форма обращения: личная, но формальная ("Добрый вечер! Меня зовут Мэри Джонс. Я работаю в газете Spingville Herald. Могу я поговорить с Биллом Ричар-дсом?"), личная, но неформальная ("Привет, это Мэри Джонс из газеты Spingville Herald. Могу я поговорить с Биллом Ричардсом?") или безличная ("Вам звонят из газеты Spingville Herald...").

В ходе предыдущих исследований эти показатели уже фиксировались — в презентациях участвовали исключительно женщины, причем все они использовали неформальный стиль обращения. Теперь необходимо определить, был ли этот выбор наилучшим.

Для испытания были отобраны 30 абонентов — 15 мужчин и 15 женщин. Они были случайным образом распределены на подгруппы, использовавшие разные стили обращения, так что в каждой из шести подгрупп, соответствовавших разным комбинациям факторов, было по пять абонентов. Стили обращения обозначены следующим образом: PF — личное формальное, PI — личное неформальное, Imp — безличное. Абоненты знали, что исследователи внимательно следят за их работой, но не знали, какой именно разговор будет прослушан. Продолжительность телефонных разговоров измерялась в секундах, прошедших от момента, когда абонент ответил на вопрос, до момента, когда он положил трубку. Результаты приведены в табл. SH.10.2.

План презентации

Структурированный Полуструктурированный Неструктурированный

38,8 41,8 32,9

42,1 36,4 36,1

45,2 39,1 39,2

34,8 28,7 29,3

48,3 36,4 41,9

37,8 36,1 31,7

41,1 35,8 35,2

43,6 33,7 38,1

УПРАЖНЕНИЯ

Таблица SH.10.2. Продолжительность телефонного разговора (в секундах) в зависимости от пола абонента и формы обращения

Форма обращения

Пол PF PI Imp

Мужской 45,6 41,7 35,3

 49,0 42,8 37,7

 41,8 40,0 41,0

 35,6 39,6 28,7

 43,4 36,0 31,8

Женский 44,1 37,9 43,3

 40,8 41,1 40,0

 46,9 35,8 43,1

 51,8 45,3 39,6

 48,5 40,2 33,2

УПРАЖНЕНИЯ

Проанализируйте данные, приведенные в табл. SH.10.2. Оцените главные эффекты каждого фактора и эффект их взаимодействия. Изложите свои выводы и рекомендации в отчете. Сформулируйте свои рекомендации, касающиеся дальнейших экспериментов.

Считаете ли вы, что длительность телефонного разговора — самый информативный результат этот эксперимента? Какие еще переменные следует исследовать? Обоснуйте свой ответ.

ПРИМЕНЕНИЕ WEB

Примените свои знания о дисперсионном анализе и оцените качество процесса расфасовки кукурузных хлопьев на заводе компании Oxford Cereals (см. главы 6,8 и 9).

После заявлений Организации потребителей, уверенных, что компания Oxford Cereal мошенничает (сокращенно — ОПУЧКОСМ), компания Oxford Cereal пожаловалась, что ее противники используют подтасованные данные. Проанализируйте ответ компании, разме-щенный на Web-странице www.prenhall.com/Springville/OC\_SelectiveData.htm), и ответьте на следующие вопросы.

1. Обоснована ли жалоба компании? Аргументируйте свой ответ.

2. Предположим, что выборки, проверенные компанией, были действительно слу-чайными. Выполните соответствующий анализ и сделайте свое заключение.

3. К каким выводам вы пришли? Чью сторону вы приняли бы, если вас пригласили стать экспертом: ОПУЧКОСМ или компании Oxford Cereal? Аргументируйте свой ответ.

СПРАВОЧНИК ПО EXCEL ГЛАВА 10 ЕН.10.1. Процедура Тьюки-Крамера

Процедура Тьюки-Крамера представляет собой двухэтапный процесс. Сначала нужно выбрать команду Сервисе Анализ данных...^Однофакторный дисперсионный анализ, чтобы вычислить выборочные средние и объемы каждой группы, количество степеней свободы в каждой группе и величину MSW. Затем из табл. Д.9 извлекается значение стьюдентизированной статистики размаха Q и создается рабочий лист, в котором для вычисления моделей разностей и критического размаха применяются обычные функции. Кроме того, на этот рабочий лист выводится сообщение о том, является ли различие между средними каждой пары групп статистически значимым.

В табл. ЕН.10.1-ЕН.10.3 показан шаблон рабочего листа, реализующего процедуру Тьюки-Крамера для сравнения прочности парашютов на основе данных, представленных на рис. 10.5. Предполагается, что процедура Анализ данных... ^Однофакторный дисперсионный анализ генерирует рабочий лист с названием Anova, на пересечении строк 5-8 и столбцов D и В содержащий выборочные средние и объемы каждой группы соответственно, в ячейке С14 — количество степеней свободы внутри групп, а в ячейке D14 — величину MSW. Шаблон также содержит в ячейке В15 стьюдентизированный размах Q (4,05), извлеченный из табл. Д.9 при уровне значимости а, равном 0,05, и заданном количестве степеней свободы. В столбце I с помощью функции ЕСЛИ выполняется сравнение абсолютных величин разностей и критических размахов, а также выводится сообщение о том, отличаются математические ожидания друг от друга или нет.

Таблица ЕН.10.3. Диапазон НЗ: 15 шаблона рабочего листа ТьюкиКрамер (строки 1 и 2 в столбцах Н и I пусты)

 Н

3 Критический

4 размах = ЕСЛИ(В15=""; "Пока статистика Q не содержится в ячейке В15; метод не корректен"; "Результаты")

S =$B$15\*G5 = ЕСЛИ(Р5>Н5;"Математические ожидания отличаются значимо "; "Математические ожидания не отличаются")

6 =$B$15\*G6 = ЕСЛИ(Р6>Н6;"Математические ожидания отличаются значимо "; "Математические ожидания не отличаются")

7 =$B$15\*G7 =ЕСЛИ(Р7>Н7;"Математические ожидания отличаются значимо "; "Математические ожидания не отличаются")

8 =$B$15\*G8 = ЕСЛИ(Р8>Н8;"Математические ожидания отличаются значимо "; "Математические ожидания не отличаются")

9 =$B$15\*G9 = ЕСПИ(Р9>Н9;"Математические ожидания отличаются значимо "; " Математические ожидания не отличаются")

10 =$B$15\*G10 =ECTIH(F10>H10;" Математические ожидания отличаются значимо "; "Математические ожидания не отличаются")

Модифицируя формулы, содержащиеся в столбцах от Е до I, и меняя выборочные средние и объемы групп в столбцах А, В и С, этот рабочий лист можно настраивать для решения аналогичных задач. Кроме того, при настройке необходимо изменить ссылки на рабочий лист ANOVA. Модифицируя рабочий лист, помните следующее.

• Извлеките объемы выборок и выборочные средние из столбцов Счет и Среднее в сводной таблице, которая начинается со строки 4.

• Извлеките количество степеней свободы в числителе и знаменателе из строк Степени свободы и Внутри группы таблицы ANOVA.

• Извлеките значение MSW из ячейки Дисперсия и строки Внутри группы таблицы ANOVA.

ЕН.10.2. Вычисление разностей между наблюдениями и медианами

Для применения критерия Левенэ сначала необходимо создать рабочий лист, вы-числяющий абсолютные величины разностей между наблюдениями и медианами каждой группы, а затем применить процедуру Сервис^Анализ данных...^Однофакторный дисперсионный анализ.

В табл. ЕН.10.4 и ЕН.10.5 продемонстрирован шаблон рабочего листа КритерийЛевенэ, вычисляющий абсолютные величины разностей между медианами групп и данными о прочности парашютов, представленными на рис. 10.5, с помощью функций МЕДИАНА hABS. Для реализации этого шаблона сначала необходимо открыть рабочий лист Данные в книге Chapter 10 . xls, скопировать метки и величины из диапазона Al: D6 в тот же диапазон нового рабочего листа КритерийЛевенэ, ввести в ячейку А8 формулу =МЕДИАНА (А2 : Аб) и скопировать эту формулу в ячейки строки вплоть до ячейки D8.

Таблица ЕН.10.4. Шаблон рабочего листа КритерийЛевенэ

 А с D

1 Поставщик 1 Поставщик 2 Поставщик 3 Поставщик 4

11:11 18,5 26,3 20,6 25,4

111!!! 24,0 25,3 25,2 19,9

1111! 17,2 24,0 20,8 22,6

5 19,9 21,2 24,7 17,5

6 18,0 24,5 22,9 20,4

7

8 = МЕДИАНА(А2:Аб) = МЕДИАНА(В2:В6) = МЕДИАНА(С2:Сб) = MEflHAHA(D2:D6)

Для того чтобы реализовать табл. ЕН.10.5, скопируйте метки из диапазона Al: D1 в диапазон Fl: II. Введите в ячейку F2 формулу =ABS (А2-А$8) и скопируйте ее в ячейки всего диапазона F2 :1 б.

ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА

1. Berenson, М. L., D. М. Levine, and М. Goldstein, Intermediate Statistical Methods and Applications: A Computer Package Approach (Englewood Cliffs, NJ: Prentice Hall, 1983).

2. Conover, W. J., Practical Nonparametric Statistics, 3rd ed. (New York: Wiley, 2000).

3. Daniel, W., Applied Nonparametric Statistics, 2nd ed. (Boston, MA: Houghton Mifflin, 1990).

4. Hicks, C. R., and К. V. Turner, Fundamental Concepts in the Design Experiments, 5th ed. (New York: Oxford University Press, 1999).

5. Kramer, C. Y., "Extension of Multiple Range Tests to Group Means with Unequal Numbers of Replications", Biometrics 12(1956): 307-310.

6. Microsoft Excel 2002 (Redmond, WA: Microsoft Corporation, 2001).

7. Montgomery, D.M., Design and Analysis of Experiments, 5th ed. (New York: John Wiley, 2001).

8. Neter, J., M. H. Kutner, C. Nachtsheim, and W. Wasserman, Applied Linear Statistical Models, 4th ed. (Homewood, IL: Irwin, 1996).

9. Tukey, J. W. "Comparing Individual Means in the Analysis of Variance", Biometrics 5 (1949): 99-114.

Предыдущая заметка [Проверка гипотез: двухвыборочные критерии](http://baguzin.ru/wp/?p=5832)

Следующая заметка

К оглавлению [Статистика для менеджеров с использованием Microsoft Excel](http://baguzin.ru/wp/?p=5285)

1. Используются материалы книги Левин и др. Статистика для менеджеров. – М.: Вильямс, 2004. – с. 681–693 [↑](#footnote-ref-1)