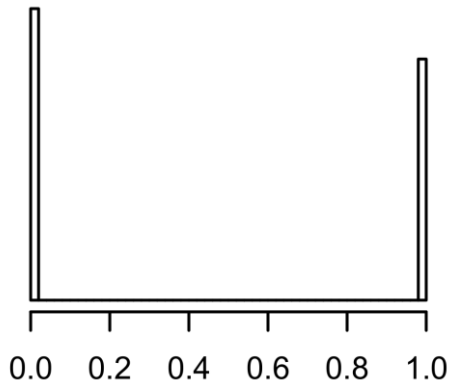


Генетика количественных признаков

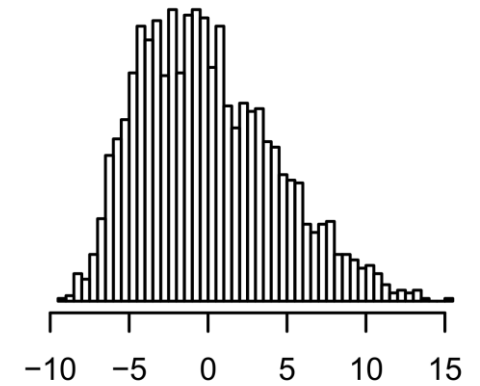
Белоногова Н.М.
к.б.н. н.с. ИЦиГ СО РАН

Качественные

Количественные



μ, σ^2



$$\sigma^2 = \sigma_G^2 + \sigma_E^2 + 2\text{cov}(G,E)$$

$$H^2 = \sigma_G^2 / \sigma^2$$

$$\sigma_G^2 = \sum \sigma_{g_i}^2 + 2\sum \sum \text{cov}(g_i g_j) + ..$$

$$h^2 = \sigma_{G(a)}^2 / \sigma^2$$

H^2, h^2 – характеристика конкретной популяции (генеральной совокупности) в конкретных условиях

не признака самого по себе (!)
не его генетического контроля

Аутизм

~90% в исследованиях близнецов

Разброс оценок от 17% до 96%

$h^2 = 54\%$, $s.e. = 5\%$

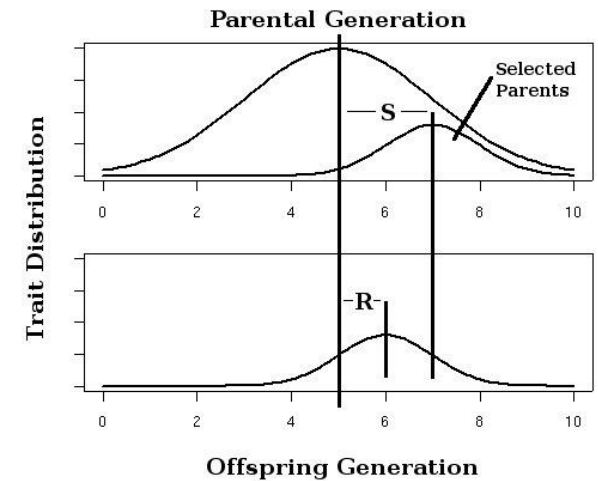
в шведской популяции (Gaugler et al., 2014)

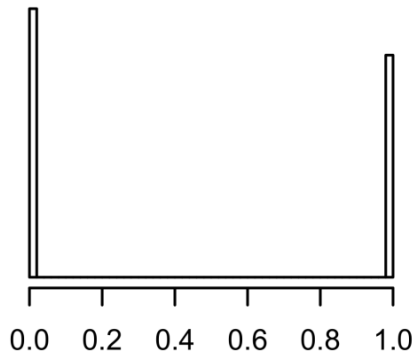
Способы оценки

- по эффекту на отбор ($h^2 = R/S$)

- исследования близнецов

- оценка случайного эффекта полигена в смешанной модели

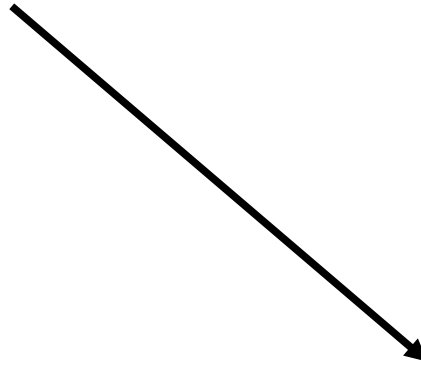




Качественные



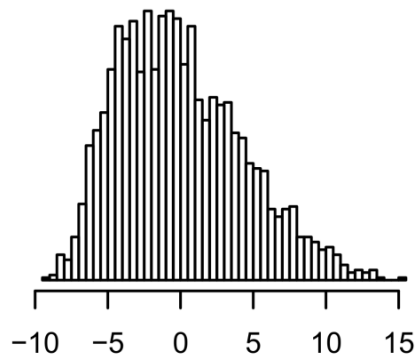
Простые

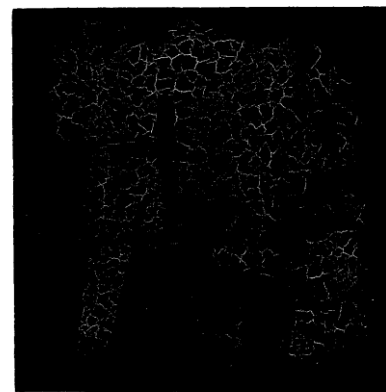


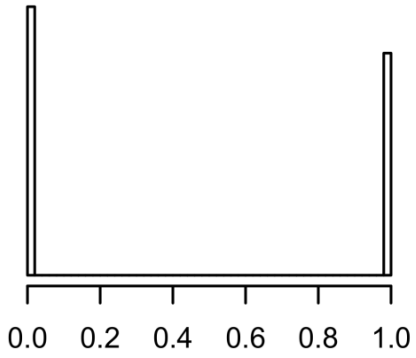
Количественные



Сложные



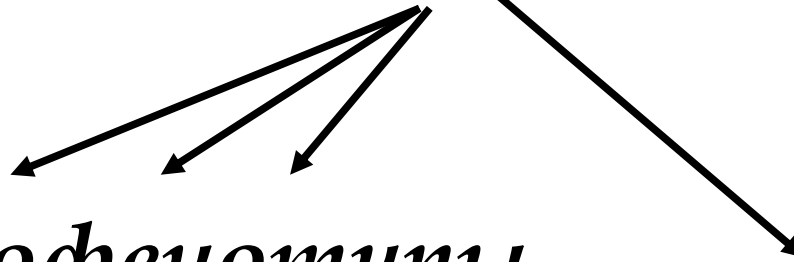




Качественные



Простые

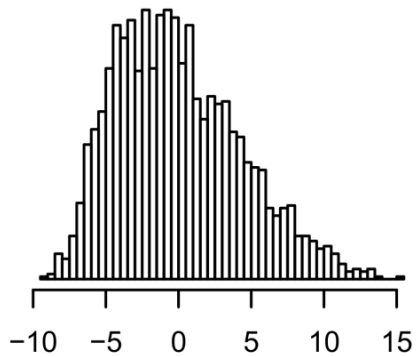


Эндотипы

Количественные



Сложные



Рост человека

2008:

~63 000 человек

54 локуса

~5% дисперсии

$$y_i = \beta_1 g_{1i} + \beta_2 g_{2i} + \dots + \beta_{54} g_{54i} + \dots$$

1886:

Гальтон,

полусумма роста родителей

~40% дисперсии

$$y_i = (y_{ip} + y_{im})/2$$

2014:

253 288 человек

$P < 5 \times 10^{-8}$: 697 SNPs ~14% дисперсии

$P < 5 \times 10^{-5}$: ~1900 SNPs ~17% дисперсии

Missing heritability

- только распространенные аллели
- эффект не-причинного варианта
- эффект оценивается с ошибкой
- в другой популяции эффект может быть другим
- игнорируется:

взаимодействие генотипов

доминирование

$$\sigma_G^2 = \sum \sigma_{g_i}^2 + 2\sum\sum \text{cov}(g_i g_j) + ..$$

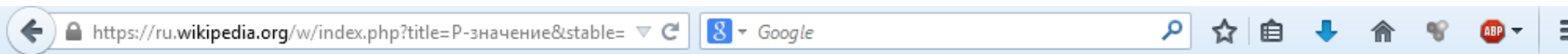
эффекты родителя и материнский эффект

взаимодействие со средой

CNV

$$\sigma^2 = \sigma_G^2 + \sigma_E^2 + 2\text{cov}(G,E)$$

$$P(D | H) \neq P(H | D)$$



P-значение

Материал из Википедии — свободной энциклопедии

[\[править\]](#) | [править вики-текст](#)

P-значение (**англ.** *P-value*) — величина, используемая при тестировании статистических гипотез. Фактически это вероятность ошибки при отклонении нулевой гипотезы (*ошибки первого рода*). Проверка гипотез с помощью P-значения является альтернативой классической процедуре проверки через критическое значение распределения.

Обычно P-значение равно вероятности того, что случайная величина с данным распределением (распределением тестовой статистики при нулевой гипотезе) примет значение, не меньшее, чем фактическое значение тестовой статистики.

Формальное определение и процедура тестирования [\[править | править вики-текст \]](#)

Пусть $t(X)$ - статистика, используемая при тестировании некоторой нулевой гипотезы H_0 . Предполагается, что если нулевая гипотеза справедлива, то распределение этой статистики известно. Обозначим функцию распределения $F(t) = P(T < t)$. P-значение чаще всего (при проверке правосторонней альтернативы) определяется как:

$$P(t) = P(T > t) = 1 - F(t)$$

При проверке левосторонней альтернативы,

$$P_0(t) = P(T < t) = F(t)$$

В случае двустороннего теста p-значение равно:

$$P(t) = 2 \min(P_0, P)$$

Если p(t) меньше заданного уровня значимости, то нулевая гипотеза отвергается в пользу альтернативной. В противном случае она не отвергается.

Преимуществом данного подхода является то, что видно при каком уровне значимости нулевая гипотеза будет отвергнута, а при каких принята, то есть виден уровень надежности статистических выводов, точнее вероятность ошибки при отвержении нулевой гипотезы. При любом уровне значимости меньше P нулевая гипотеза отвергается, а при больших значениях - нет.

- Заглавная страница
- Рубрикация
- Указатель А—Я
- Избранные статьи
- Случайная статья
- Текущие события

Участие

- Сообщить об ошибке
- Портал сообщества
- Форум
- Свежие правки
- Новые страницы
- Справка
- Пожертвования

Инструменты

- Ссылки сюда
- Связанные правки
- Спецстраницы
- Постоянная ссылка
- Сведения о странице
- Цитировать страницу

$$P(D | H) \neq P(H | D)$$

Электронный учебник - ...

www.statsoft.ru/home/textbook/default.htm

Google

Что такое статистическая значимость (p-уровень)? Статистическая значимость результата представляет собой оцененную меру уверенности в его "истинности" (в смысле "репрезентативности выборки"). Выражаясь более технически, p-уровень (этот термин был впервые использован в работе Brownlee, 1960) это показатель, находящийся в убывающей зависимости от надежности результата. Более высокий p-уровень соответствует более низкому уровню доверия к найденной в выборке зависимости между переменными. Именно, p-уровень представляет собой вероятность ошибки, связанной с распространением наблюдаемого результата на всю популяцию. Например, p-уровень = .05 (т.е. 1/20) показывает, что имеется 5% вероятность, что найденная в выборке связь между переменными является лишь случайной особенностью данной выборки. Иными словами, если данная зависимость в популяции отсутствует, а вы многократно проводили бы подобные эксперименты, то примерно в одном из двадцати повторений эксперимента можно было бы ожидать такой же или более сильной зависимости между переменными. (Отметим, что это не то же самое, что утверждать о заведомом наличии зависимости между переменными, которая в среднем может быть воспроизведена в 5% или 95% случаев; когда между переменными популяции существует зависимость, вероятность повторения результатов исследования, показывающих наличие этой зависимости называется статистической мощностью плана. Подробнее об этом см. в разделе [Анализ мощности](#)). Во многих исследованиях p-уровень .05 рассматривается как "приемлемая граница" уровня ошибки.

[В начало](#)

Как определить, является ли результат действительно значимым. Не существует никакого способа

Copyright (c) StatSoft, Inc. 1984-2014. Статистический учебник является общедоступным сервисом, предоставленным компанией StatSoft, Inc., создателем системы STATISTICA - высокопроизводительного программного обеспечения для графического и статистического анализа.