

# ОПТИМИЗАЦИЯ ФЕРМЕНТАЦИОННЫХ СРЕД

## ОСНОВНЫЕ ПОНЯТИЯ



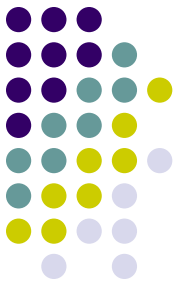
# существуют среды, используемые при решении разных задач:



- генетических исследований, при обработке штаммов мутагенами;
- селекции мутантов — полноценные и селективные среды;
- длительного хранения и пересевов штаммов микроорганизмов;
- оптимизации продуктивности штамма в колбах;
- для инокулята и посевного материала (в том числе вегетативного и спорового);
- оптимизации продуктивности в ферментерах производственного масштаба.

# Выбор критерия оптимизации

(применительно к средам в колбах)



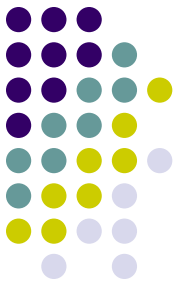
- количество целевого продукта;
- производительность по целевому продукту;
- выход целевого продукта по субстрату;
- минимизация стоимости среды для получения единицы целевого продукта.

# Выбор исходных компонентов среды



- При подборе сред в первую очередь следует обратить внимание на химический состав биомассы микроорганизмов и на химический состав внеклеточного продукта, если он является целевым.
- Основные компоненты среды — источники углерода (и одновременно энергии), азота, фосфора, серы, микроэлементов, ростовых факторов и витаминов для начала роста биомассы.

# Определение соотношения компонентов среды

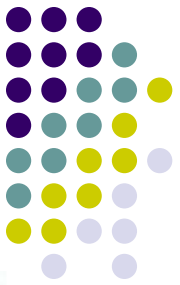


- . Рецептуры сред для промышленных штаммов охраняются как большой секрет, и многие микробиологи занимаются их подбором.
- Когда в составе среды только один компонент не изучен и необходимо определить его оптимальную концентрацию, достаточно провести серию опытов, в которых исследуемый компонент будет находиться в разных концентрациях.
- Проведения опытов с 4—5 разными концентрациями (уровнями) компонента достаточно, чтобы получить представление о характере его влияния.

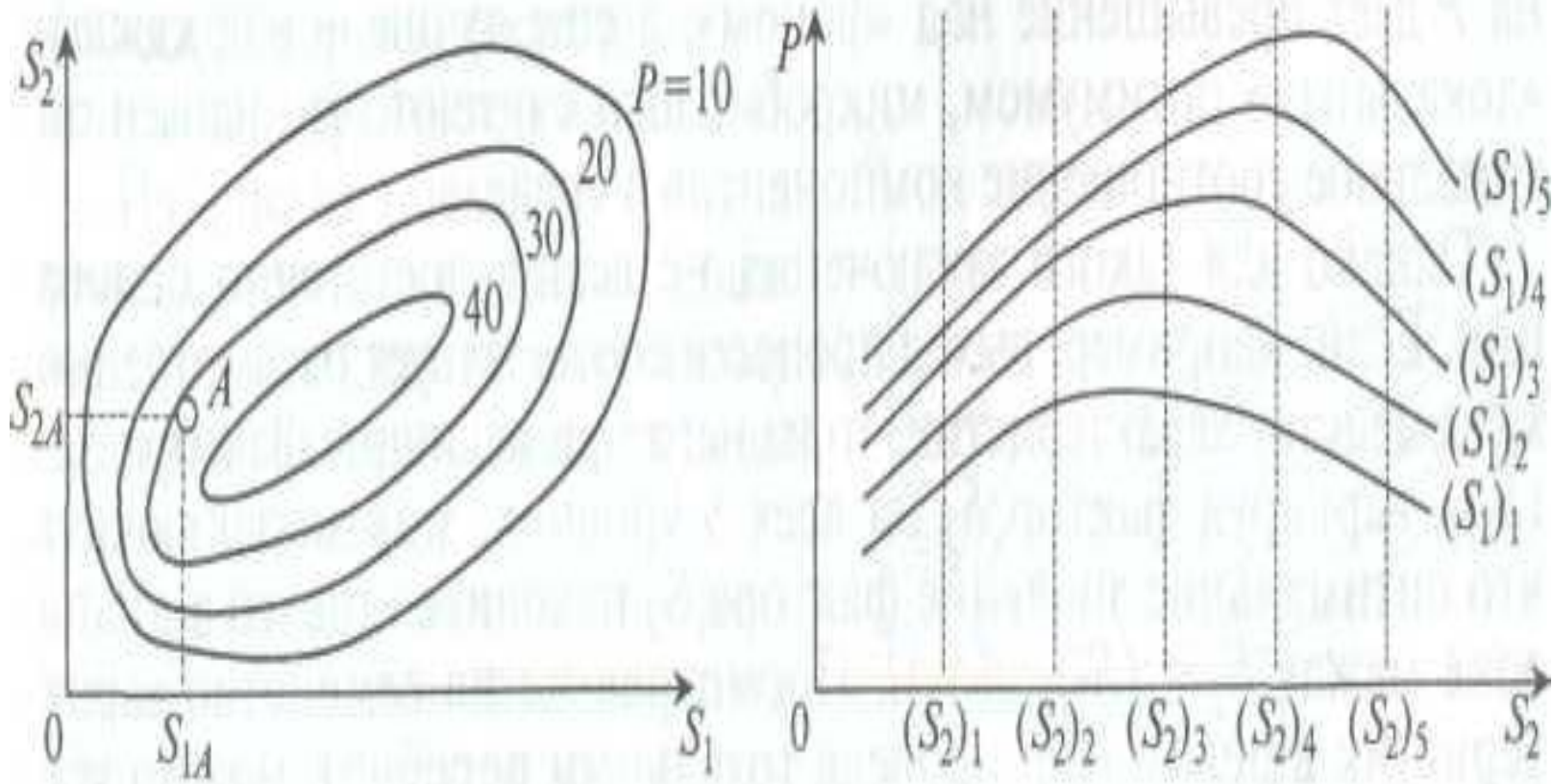
# ТРАДИЦИОННЫЕ МЕТОДЫ ИЗУЧЕНИЯ МНОГОФАКТОРНЫХ ЗАВИСИМОСТЕЙ



- Необходимо построить *поверхность отклика*, напоминающую топографическую карту. На этой карте на двух осях координат отложена величина факторов  $S_1$  и  $S_2$ , а в самом графике проведены контурные *линии равного уровня*, соответствующие одинаковому выходу  $P$  — параметру оптимизации.



для построения «топографической карты»  
нужно «изрешетить» всю площадь в  
изучаемом диапазоне  $S_1$  и  $S_2$  опытами.

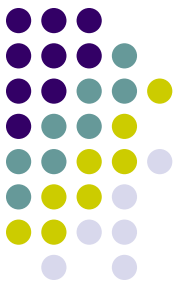


# Отдельное изучение каждого фактора.



- при подборе сред микробиологи применяют исходный «фон» — определенное соотношение уровней факторов (концентраций компонентов) в среде. На этом «фоне» ставят однофакторные эксперименты по каждому из  $n$  факторов и получают столько же кривых, сколько изучают факторов.





# Метод Гаусса — Зайделя

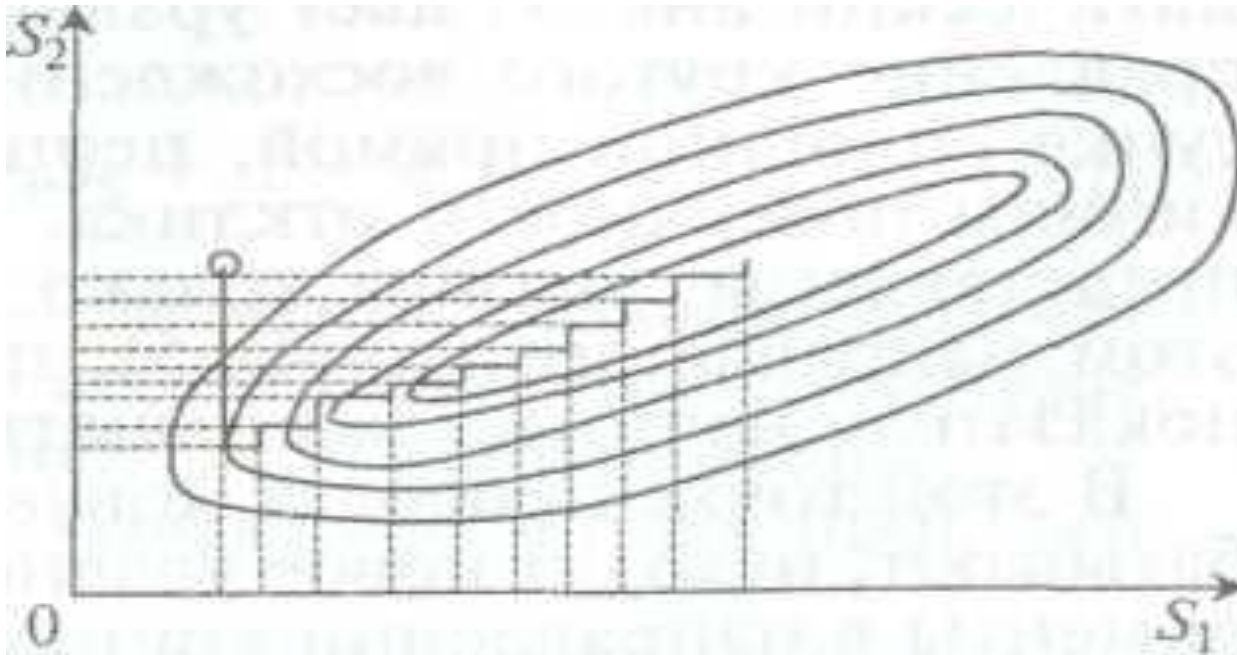
- этот метод называют «последовательным изучением каждого фактора». Здесь частные зависимости  $P(S_1)$ ,  $P(S_2)$ ,  $P(S_n)$  находят не сразу на одном и том же «фоне», а по очереди. Сначала определяют  $P(S_i)$ , анализируют эту зависимость, находят частный оптимум для фактора  $S_i$ , затем меняют «фон», установив в нем значение  $S_i$ , равное этому частному оптимуму.



# Недостатки метода:

- метод Гаусса — Зайделя следует повторять пока частные оптимумы по отдельным факторам не перестанут изменяться от одной серии опытов к другой.
- метод имеет и еще один недостаток по сравнению с предыдущими: необходимо *поочередное* выполнение серий экспериментов, так как изменение «фона» требует анализа результатов каждой серии.

# Процедура оптимизации по методу Гаусса-Зайделя

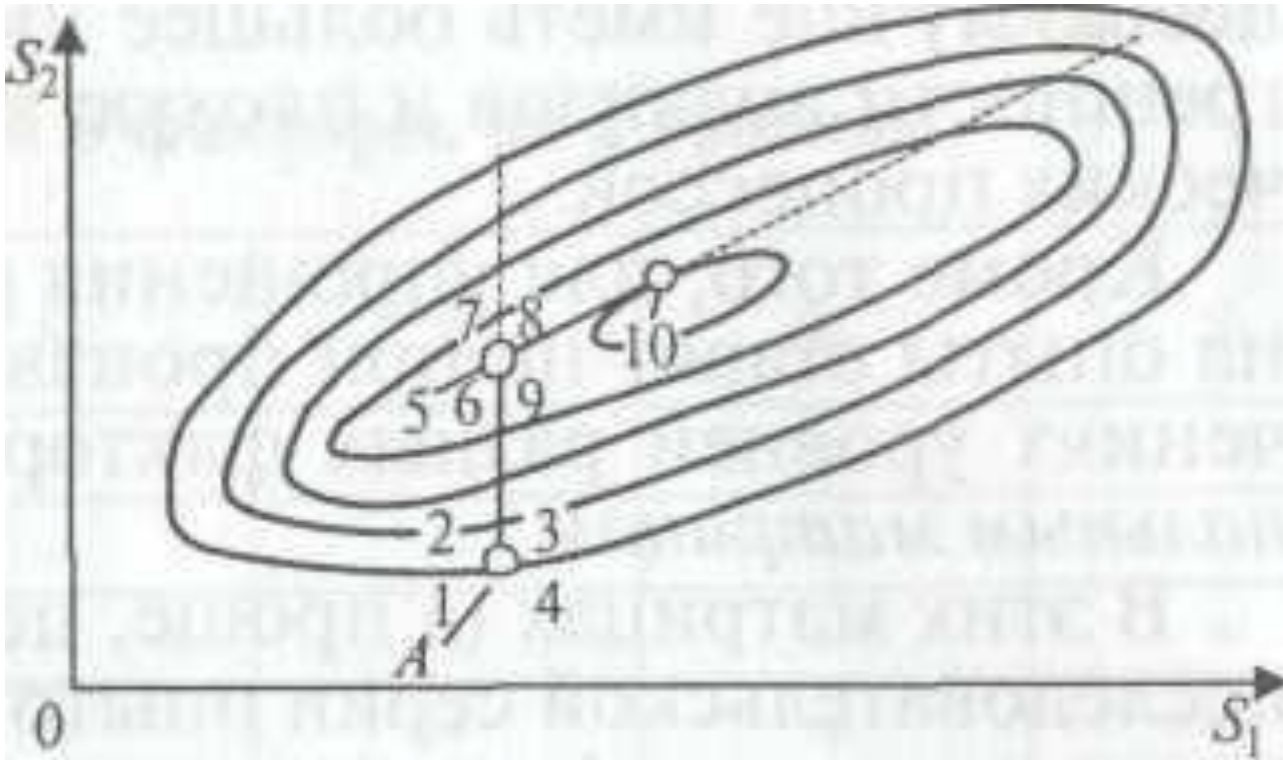


# МЕТОД БОКСА — УИЛСОНА



- По этому методу вблизи исходной точки («фона») ставится специальным образом спланированная небольшая серия опытов, в которой одновременно варьируются все изучаемые факторы, каждый на 2 уровнях (верхнем и нижнем).
- Результаты этих опытов *математически обрабатывают* для получения приближенного математического описания процесса в этой локальной области. Для двух уровней варьирования факторов находят линейное уравнение (*уравнением регрессии*), величина факторов входит в первой степени.

# Процедура оптимизации по методу Бокса-Уилсона



# Исследовательская серия опытов



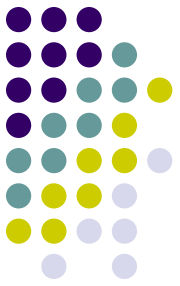
- сначала ставится небольшая серия опытов на двух уровнях — верхнем и нижнем. Для каждого фактора эти уровни отличаются от *основного уровня* (исходного уровня «фона») на одну и ту же величину.
- Эта величина называется *интервалом варьирования* и обозначается  $X$ . Для разных факторов величина  $\lambda$  может быть разной.

Для чего ставится исследовательская серия опытов?

Чтобы получить линейное уравнение, связывающее выходной параметр

оптимизации  $P$  с влияющими факторами:

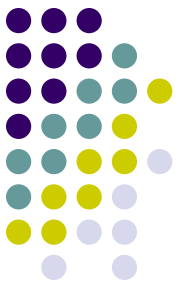
$$P = b_0 + b_1 S_1 + b_2 S_2 + b_3 S_3 + \dots + b_n S_n.$$





- для упрощения расчетов по методу Бокса—Уилсона опыты ставят не при произвольным образом измененных значениях уровней разных факторов, а по так называемым *ортогональным матрицам*
- В этих матрицах (перечне вариантов, взятых в данной исследовательской серии опытов) из всех возможных  $2^n$  вариантов опытов, когда  $n$  факторов варьируется на 2 уровнях, выбирают небольшое количество вариантов со следующими свойствами





## Свойства матрицы:

- 1. в каждой серии количество вариантов опытов с верхним уровнем *каждого фактора* равно количеству вариантов с нижним уровнем того же фактора;
- 2. верхний уровень *любого фактора* сочетается одинаковое число раз и с верхними, и с нижними уровнями *всех остальных факторов*, он как бы проверяется на усредненном фоне, в котором влияние уровней остальных факторов нивелируется. Это же положение справедливо и для нижних уровней.

- Планы ортогональных матриц разработаны для различного числа факторов  $n$  и могут включать разное число вариантов планирования. Практически в исследовательских сериях стараются уменьшить число вариантов опытов. Для этого целесообразно выбрать ближайшую матрицу с числом вариантов большим, чем  $(n + 2)$ .



**Планы экспериментов имеются в различных справочниках и руководствах.**



**Матрица планирования для 2 факторов на 2 уровнях**

№ варианта	$S_1$	$S_2$
1	-	-
2	+	-
3	-	+
4	+	+

Матрица планирования для 6 факторов на 2 уровнях состоящая из 8 экспериментов, может быть использована также для 3 факторов (это — полный факторный эксперимент), для 4 и для 5 факторов (вычеркивая за ненадобностью 1, 2 или 3 столбца из этой матрицы).



Матрица планирования для 6 факторов на 2 уровнях

№ вариантов	$S_1$	$S_2$	$S_3$	$S_4$	$S_5$	$S_6$
1	-	-	-	-	-	-
2	+	-	-	+	+	-
3	-	+	-	+	-	+
4	+	+	-	-	+	+
5	-	-	+	+	+	+
6	+	-	+	-	-	+
7	-	+	+	-	+	-
8	+	+	+	+	-	-



- После выбора матрицы планирования, числа варьируемых факторов и интервалов планирования можно переходить от условных обозначений матриц, выраженных чередованием знаков (-) и (+), к их физическому наполнению реальными значениями концентраций компонентов среды, которые предстоит проверять в опытах при оптимизации состава среды. Все результаты удобно записывать в сводную таблицу.

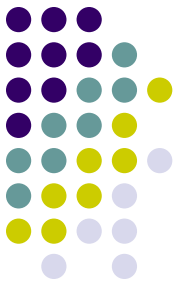
# Пример оптимизации 4-х компонентной среды по методу Бокса-Уилсона



Наименование этапов работы	Наименование факторов (концентраций компонентов среды)				Конечная концентрация антибиотика, ед/мл						
	$S_1$ , крахмал, %	$S_2$ , $(NH_4)_2SO_4$ , %	$S_3$ , кукурузный экстракт, %	$S_4$ , соевая мука, %	$P_1$	$P_2$	$\bar{P}$	$(P_j - \bar{P})$	$(P_j - \bar{P})^2$	$\hat{P}$	
Основной уровень ( $S_i$ ) <sub>0</sub>	6,0	0,6	0,75	2,0	—	—	—	—	—	—	
Максимальный уровень ( $S_i$ ) <sub>max</sub>	10,0	2,0	2,0	5,0	—	—	—	—	—	—	
Минимальный уровень ( $S_i$ ) <sub>min</sub>	0,0	0,0	0,0	0,0	—	—	—	—	—	—	
Интервал варьирования $\lambda_i$	0,7	0,1	0,15	0,3	—	—	—	—	—	—	
Матрица планирования эксперимента	№ варианта										
	1	(-)5,3	(-)0,5	(-)0,6	(-)1,7	5600	5800	5700	100	10000	5553
	2	(+)6,7	(-)0,5	(-)0,6	(+)2,3	7200	7000	7100	100	10000	7149
	3	(-)5,3	(+)0,7	(-)0,6	(+)2,3	6750	6450	6600	150	22500	6663
	4	(+)6,7	(+)0,7	(-)0,6	(-)1,7	6030	6210	6120	90	8100	6179
	5	(-)5,3	(-)0,5	(+)0,9	(+)2,3	6930	7030	6970	60	3600	7027
	6	(+)6,7	(-)0,5	(+)0,9	(-)1,7	6740	6400	6520	120	14400	6563
	7	(-)5,3	(+)0,7	(+)0,9	(-)1,7	6080	5940	6040	70	4900	6057
8	(+)6,7	(+)0,7	(+)0,9	(+)2,3	8000	7600	7800	200	40000	7653	
$\Sigma(P - \bar{P})^2$	—	—	—	—	—	—	—	—	113500	—	

Продолжение таблицы

# ОПТИМИЗАЦИЯ 4-Х КОМПОНЕНТНОЙ СРЕДЫ ПО МЕТОДУ БОКСА-УИЛСОНА



Наименование этапов работы		Наименование факторов (концентраций компонентов среды)				Конечная концентрация антибиотика, ед/мл					
		$S_1$ , крахмал, %	$S_2$ , (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> , %	$S_3$ , куку- рузный экс- тракт, %	$S_4$ , соевая мука, %	$P_1$	$P_2$	$\bar{P}$	$(P_j - \bar{P})$	$(P_j - \bar{P})^2$	$\hat{P}$
Коэффициенты регрессии $b_i$		+283	+30	+222	+515	$(b_0 = 6603)$			—	—	—
Произведение $b_i \lambda_i$		+198	+3	+33	+155	$\Sigma \Sigma (P_j - \bar{P})^2 = 2 \cdot 113500 = 227000$					
Запас $\Delta_i$		4,0	1,4	1,25	3,0	$\Sigma (\bar{P} - \hat{P})^2 = 58258$					
Критичность $ \Delta_i / (b_i \lambda_i) $		0,02	0,47	0,038	0,019						
Шаг $\Delta S_i$		+0,64	—	+0,11	+0,5	Расч. знач. $\hat{P}$	Экспер. знач. $P$	—	—	—	—
Опыты кругого восхождения	9	6,6	0,6	0,86	2,5	7879	7140	—	—	—	—
	10	7,3	0,6	0,97	3,0	9155	7630	—	—	—	—
	11	7,9	0,6	1,08	3,5	10430	8350	—	—	—	—
	12	8,6	0,6	1,19	4,0	11706	8510	—	—	—	—
	13	9,2	0,6	1,30	4,5	12982	6220	—	—	—	—

# оптимизация среды по методу Бокса-Уилсона



- После выбора матрицы планирования, числа варьируемых факторов и интервалов планирования можно переходить от условных обозначений матриц, выраженных чередованием знаков (-) и (+), к их физическому наполнению реальными значениями концентраций компонентов среды, которые предстоит проверять в опытах при оптимизации состава среды.



# МАТЕМАТИЧЕСКИЕ ПРОЦЕДУРЫ В МЕТОДЕ БОКСА—УИЛСОНА



- основной уровень («фон»), интервалы варьирования, минимальный и максимальный уровни фактора, при которых исследователь допускает протекание процесса, определяются интуицией.
- необходимо найти физические значения верхнего и нижнего уровней факторов.
- Для этого к основному уровню нужно прибавить интервал варьирования — для верхнего уровня или вычесть его — для нижнего уровня.

# МАТЕМАТИЧЕСКИЕ ПРОЦЕДУРЫ В МЕТОДЕ БОКСА—УИЛСОНА



- кроме варьируемых факторов в составе среды могут быть факторы, значение которых одинаково для всех вариантов среды.
- надо не забывать при приготовлении сред добавлять и эти постоянные для всех сред компоненты, хотя в матрице планирования они и не указаны.

# Постановка эксперимента



- В микробиологических исследованиях принято для надежности повторять опыты и не по одному разу. Это касается и опытов с планированием эксперимента при подборе сред.
- в данных опытах есть одна особенность:
- - не интересны конкретные цифры для каждой отдельной среды:
- - *вся серия в целом* служит одной задаче: найти более точно описание всей области эксперимента уравнением, а с ним — и более точное направление движения к оптимуму;
- - каждая точка матрицы помогает другой и они дают *совокупный результат*.

**Это позволяет снизить требования к повторению опытов.**

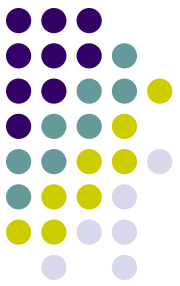
# Расчет коэффициентов уравнения регрессии



- Свободный член уравнения  $b_0$  находится как среднее из значений  $P$  для всех  $N$  вариантов матрицы:

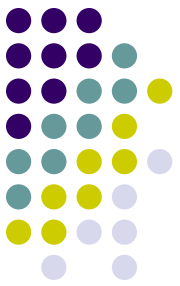
$$b_0 = \frac{\sum_{u=1}^N \bar{P}_u}{N},$$

где  $\bar{P}_u$  — среднее значение величины  $P$  для всех повторений для  $u$ -го варианта матрицы.



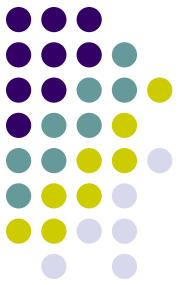
$$b_1 = \frac{\sum_{u=1}^N (\bar{P}_u S_{iu+}) - \sum_{u=1}^N (\bar{P}_u S_{iu-})}{N}$$

- Коэффициент регрессии  $i$ -го фактора определяется как разность сумм выходов  $\bar{P}_u$ , в которых фактор  $S_{i^+}$  находился со знаком (+), и выходов  $\bar{P}_u$ , где фактор  $S_{i^-}$  находился со знаком (-), деленная на число вариантов в матрице планирования  $N$ :



- Если принять кодированные значения факторов  $X_i$  т.е.  $S_{iu}$ , как (+1) для верхнего уровня и (-1) для нижнего, то получаем:

$$b_i = \frac{\sum_{u=1}^N (\bar{P}_u S_{iu})}{N}.$$



- Например, для фактора  $S_2$  в таблице коэффициент регрессии вычисляют следующим образом:

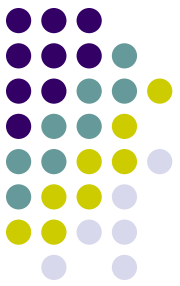
$$b_2 = \frac{(-1) \cdot 5700 + (-1) \cdot 7100 + 1 \cdot 6600 + 1 \cdot 6120 + (-1) \cdot 6970}{8} + \frac{(-1) \cdot 6520 + 1 \cdot 6040 + 1 \cdot 7800}{8} = +30.$$



- Рассчитанные таким образом коэффициенты имеют разные знаки. Знак (+) означает, что при увеличении данного фактора происходит возрастание параметра оптимизации  $P$ , знак (-) - наоборот, уменьшение  $P$ .
- Чем больше коэффициент регрессии, тем больше он влияет на результат процесса.



На основе коэффициентов регрессии  
можно записать уравнение



$$\bar{P} = 6603 + 283 S_1 + 30 S_2 + 222 S_3 + 515 S_4$$

Чтобы перейти к натуральным переменным, нужно подставить  
выражение для кодированных факторов:

$$S_i = \frac{S_i - S_{0i}}{\lambda_i},$$

где  $S_i$  и  $S_{0i}$  — натуральные значения уровня  $i$ -го фактора в данном опыте и основного уровня этого фактора;  $\lambda_i$  — интервал варьирования  $i$ -го фактора в натуральных величинах

# Определение запаса для движения в направлении крутого восхождения



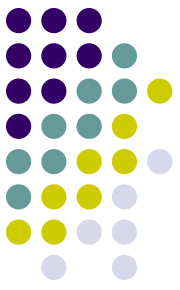
- При определении программы крутого восхождения от основного уровня необходимо знать для каждого фактора запас  $\Delta_i$ , - в сторону движения [для коэффициентов регрессии со знаком (+) - в сторону увеличения, со знаком (-) - в сторону уменьшения]:

$$\Delta_i = S_{i \max} - S_{0i}, \text{ если } b_i > 0 \text{ или } S_{0i} - S_{i \min}, \text{ если } b_i < 0. (7.6)$$

# Определение вспомогательного показателя «критичность».

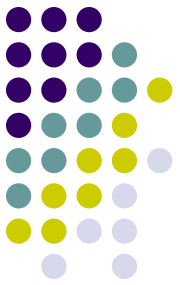


- Этот показатель представляет собой модуль отношения запаса  $\Delta_i$ , и произведения  $b_i \lambda_i$ , - и физически выражает в сравнительном аспекте, сколько сможет уместиться в «запасе» шагов в направлении крутого восхождения для разных факторов с учетом их степени влияния на результат процесса.
- Чем ниже значение показателя критичности для фактора, тем меньшее (по сравнению с другими) число шагов уместится в «запасе» для этого фактора.



## Выбор шага крутого восхождения

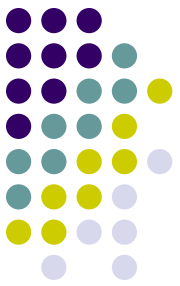
- Обычно крутое восхождение проводится путем равномерного пошагового приращения в каждом последующем опыте величины уровня фактора.
- При этом для одного из факторов — наиболее критичного — величина приращения (шага) выбирается, а для всех остальных — рассчитывается так, чтобы их значения были пропорциональны произведениям  $b_i \lambda_i$ .



# Выбор шага крутого восхождения

- Обычно выбирают относительно немного шагов в направлении крутого восхождения: 5—8, не более.
- По наиболее критичному фактору выбранное количество шагов  $m$  позволяет определить величину шага  $\Delta S_i$

$$\Delta S_i = \frac{\Delta_i}{m}$$



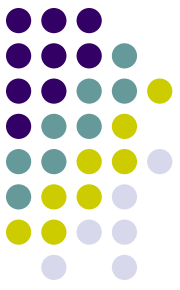
# Выбор шага крутого восхождения

- Если величина базового шага равна 0,5, то пропорционально произведениям  $b_i \lambda_i$ , - найдем величину шага  $\Delta S_i$ , для факторов  $S_i$ :

$$\Delta S_i = \frac{\Delta S_6}{(b\lambda)_6} (b\lambda)_i;$$

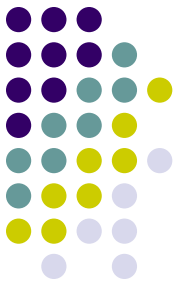
$$\Delta S_1 = \Delta S_4 \frac{b_1 \lambda_1}{b_4 \lambda_4} = 0,5 \frac{198}{155} = 0,64;$$

$$\Delta S_3 = \Delta S_4 \frac{b_3 \lambda_3}{b_4 \lambda_4} = 0,5 \frac{33}{155} = 0,11.$$



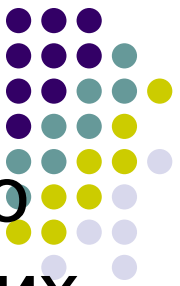
- Для фактора проставлен прочерк.
- Значит по результатам анализа установлено, что данный фактор не оказывает существенного влияния на результат процесса, а потому его значение нет смысла менять.
- Как определить, что тот или иной фактор является значимым или незначимым?
- Для этого в планировании эксперимента предусмотрена специальная процедура.

# СТАТИСТИЧЕСКАЯ ОЦЕНКА РЕЗУЛЬТАТОВ



- В биологических процессах результат процесса обычно неоднозначен.
- Существует какой-то уровень колебаний, возможных по случайным причинам.
- Природа этих колебаний (источников неоднородности) может быть различной. Это ошибки в определении результата, и различия в значении факторов при проведении эксперимента, и неточности приготовления сред, и вообще присущая биологическим объектам неоднородность.





- коэффициенты регрессии определяют по значениям выходного показателя, для них есть уровень случайных изменений, который находят с помощью статистической обработки.
- Уровень воспроизводимости процесса характеризуется *дисперсией*.
- Для определения дисперсии воспроизводимости существуют два способа.

# Первый способ— многократное повторение опытов для одного и того же варианта (среды).



- В этом случае имеем ряд значений  $P$ , их количество обозначим  $\gamma$ , оно не должно быть меньше 8—10.
- Нужно найти среднюю величину  $\bar{P}$  для всех  $\gamma$  опытов, все отклонения от среднего  $(P - \bar{P})$ , квадраты этих отклонений и сумму квадратов  $\sum (P_j - \bar{P})^2$  для всех  $\gamma$  повторений.

**Дисперсию воспроизводимости определяют по формуле:**



$$\sigma^2 = \frac{\sum_{j=1}^{\gamma} (P_j - \bar{P})^2}{\gamma - 1}$$

Это не среднеквадратичное отклонение, а дисперсия воспроизводимости каждого отдельного измерения.



- Если выполнить несколько серий опытов на одной и той же среде (в каждой серии число опытов  $\gamma$ ), затем для каждой серии рассчитать свою среднюю величину  $P$ , и посмотреть, какова будет дисперсия воспроизводимости для этих *средних значений*, то получим:

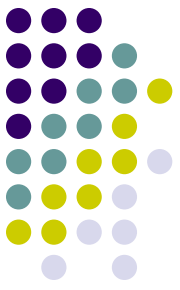
$$\sigma_{\bar{P}}^2 = \sigma^2 / \gamma.$$

# Второй способ— расчет дисперсии процесса по данным повторных экспериментов матрицы планирования.

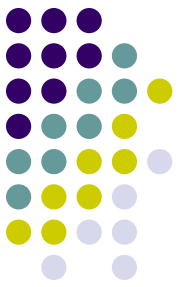


- Он более удобен при планировании эксперимента, так как повторения в каждом варианте матрицы предусмотрены. Если  $\gamma$  — число повторений в каждом варианте среды, а  $N$  — общее число вариантов сред в матрице, то дисперсию воспроизводимости *единичного значения* находят по формуле

$$\sigma^2 = \frac{\sum_{u=1}^N \sum_{j=1}^{\gamma} (P_j - \bar{P})^2}{(\gamma - 1) N}.$$

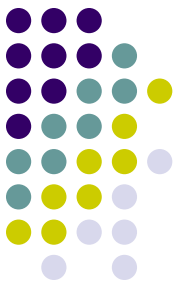


- Надежность вычисления дисперсии воспроизводимости определяется количеством «лишних» опытов, необходимых для нахождения дисперсии.
- Количество «лишних» опытов называют *числом степеней свободы  $f$* . Для первого способа оно равно  $(\gamma - 1)$ , для второго —  $N(\gamma - 1)$ . Эта величина является важным статистическим показателем.



- для надежного определения дисперсии воспроизводимости процесса число степеней свободы  $f$  должно быть не менее 5—8.
- Чем больше дисперсия, тем хуже воспроизводимость процесса.
- Корень квадратный из дисперсии воспроизводимости процесса называют *стандартным отклонением*, или *стандартной ошибкой*.

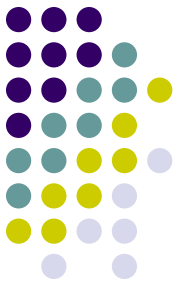
# Определение значимости коэффициентов регрессии.



- Уровень возможных случайных колебаний коэффициентов регрессии называется *доверительным интервалом*, обозначается буквой  $\varepsilon$  и его вычисляют по формуле:

$$\varepsilon = t \sqrt{\frac{\sigma^2}{N\gamma}} = t \sqrt{\frac{\sum \sum (P_u - \bar{P})^2}{(\gamma - 1) N^2 \gamma}}.$$





- В этой формуле появляется новый коэффициент  $t$  — критерий Стьюдента, определяемый по таблицам.
- Для надежности оценки до 95% критерий Стьюдента зависит только от числа степеней свободы  $f$  при которых находим дисперсию воспроизводимости.

значения критерия Стьюдента для наиболее часто встречающихся значений  $f$  при  $P = 0,95$ :



$f$	$t$	$f$	$t$
1	12,71	10	2,23
2	4,30	12	2,18
3	3,18	16	2,12
4	2,78	20	2,09
5	2,57	30	2,04
6	2,45	60	2,00
7	2,37	100	1,98
8	2,31		

вначале критерий  $t$  резко уменьшается с возрастанием  $f$ , а затем его значение стабилизируется.

- Доверительный интервал  $\varepsilon$  имеет одно и то же значение для всех коэффициентов  $b_i$  (в кодированном виде).
- Сравнение коэффициента регрессии с доверительным интервалом и позволяет сделать вывод о его значимости.
- Все коэффициенты  $b_i$ , значения которых ниже доверительного интервала, незначимы. Их можно считать нулевыми.
- Если величина коэффициента больше  $\varepsilon$ , он значим. Отрицательные коэффициенты регрессии тоже значимы, если они по модулю превосходят  $\varepsilon$ .

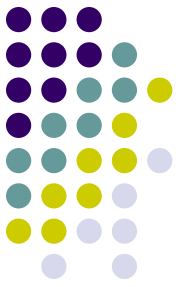


## **Незначимость коэффициентов может быть вызвана различными причинами:**



- 1. взяты слишком малые интервалы варьирования фактора;
- 2. плохая воспроизводимость процесса — все различия в выходе нивелируются ошибкой опыта;
- 3. данный фактор находится на уровне, близком к оптимальному;
- 4. данный фактор не влияет на процесс вообще, по крайней мере в изученной области.

# Адекватность математического описания процесса



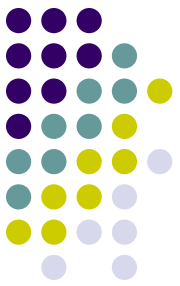
- Кроме оценки значимости коэффициентов в процедуре статистического анализа предусмотрена оценка адекватности полученного математического описания в целом.
- Для этого сначала находят *дисперсию адекватности*, которая характеризует отклонение рассчитанных по уравнению значений выходного показателя  $P$  от найденного в эксперименте  $\bar{P}$ .

для варианта 6 величину  $P$  можно  
определить так



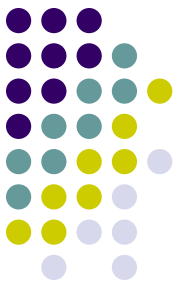
$$\begin{aligned}\hat{P} &= b_0 + b_1 - b_2 + b_3 - b_4 = \\ &= 6603 + 283 - 30 + 222 - 515 = 6563.\end{aligned}$$

# Дисперсия адекватности определяется по формуле



$$\sigma_a^2 = \frac{\sum \left( \bar{P}_u - \hat{P}_u \right)^2}{N - n - 1},$$

где  $n$  — число факторов в уравнении;  $N$  — число вариантов опытов (условий), по которым определяется дисперсия адекватности;  $u$  — номер варианта среды, к которому относятся  $\bar{P}$  и  $P$ .



- Знаменатель этой дроби представляет собой *число степеней свободы*  $f_a$  дисперсии адекватности:
- $f_a = N - n - 1$ ;
- т.е. количество «лишних» опытов, имеющихся в плане сверх минимально необходимых  $(n + 1)$  — по числу коэффициентов в уравнении.
- Степень адекватности математического описания оценивают по критерию Фишера  $F$ , который вычисляют по формуле:

$$F = \sigma_a^2 \gamma / \sigma_P^2.$$





- Чтобы найти адекватность уравнения, необходимо критерий Фишера сравнить с табличным  $F_T$ , имеющимся в справочниках по статистике, также для надежности оценки 95%.
- Уравнение считается адекватным, если  $F < F_T$ , и наоборот.
- $F_T$  в таблице определяют исходя из двух видов степеней свободы:  $fr$  — для воспроизводимости самого процесса и  $fa$  — для дисперсии адекватности.

# Расчет программы крутого восхождения



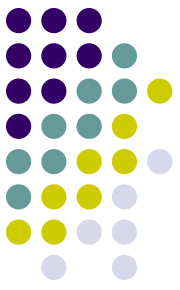
- Для данного примера рассчитывают условия для 6 шагов крутого восхождения, т. е. 6 новых сред.
- На каждом шаге к предыдущему уровню фактора прибавляется или от него отнимается рассчитанное значение шага.
- По наиболее критичному фактору последний шаг будет совпадать с максимальным или минимальным его уровнем, по другим факторам — несколько не доходить до них.
- Составляют 6 сред крутого восхождения и воспроизводят процесс с этими рассчитанными средами. Для факторов с отрицательным значением коэффициента регрессии при каждом шаге *отнимается*, а не прибавляется величина рассчитанного  $\Delta S$ ;

# Анализ результатов крутого восхождения



- сначала нужно все значения уровней факторов пересчитать в кодированные.
- затем рассчитать величину  $P$  аналогично тому, как это делалось для опытов в исследовательской матрице, подставив вместо кодированных значений  $S_i$  (-1) и (+1) значения, вычисленные по формуле:

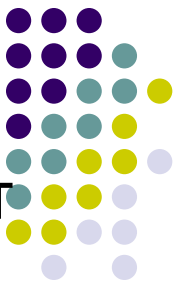
$$S_i = \frac{\tilde{S}_i - \tilde{S}_{i0}}{\lambda_i},$$



- В колонке рядом с расчетными данными указать экспериментальные.
- Часто столь высоких значений выхода, как ожидали, не получают.
- Рост выходного показателя может идти не все время — после 4-го шага ( 12-го опыта) начинается даже падение выхода.
- Значит линейное приближение недостаточно. Но направление движения к оптимуму найдено.
- Далее нужно поставить новую матрицу планирования с центром в новой точке и наметить программу крутого восхождения.

- расчет по этому методу необходимо повторять до тех пор, пока линейное уравнение не станет неадекватным. В этом случае надо использовать другие методы, описывающие процесс уже *уравнением второго порядка*:

$$P = b_0 + \sum_{i=1}^n (b_i S_i) + \sum_{i,j=1}^n (b_{ij} S_i S_j) + \sum_{i=1}^n (b_{ii} S_i^2).$$



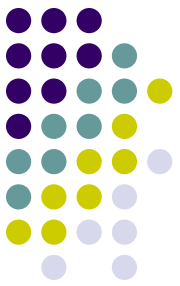
# МНОГОУРОВНЕВЫЕ ПЛАНЫ ЭКСПЕРИМЕНТА



- методом является *аддитивно-решетчатое описание* процесса.
- В этом случае объект описывается *аддитивно-нелинейным решетчатым уравнением*:
- 

$$P = b_0 + f_1[S_1] + f_2[S_2] + \dots + f_n[S_n].$$

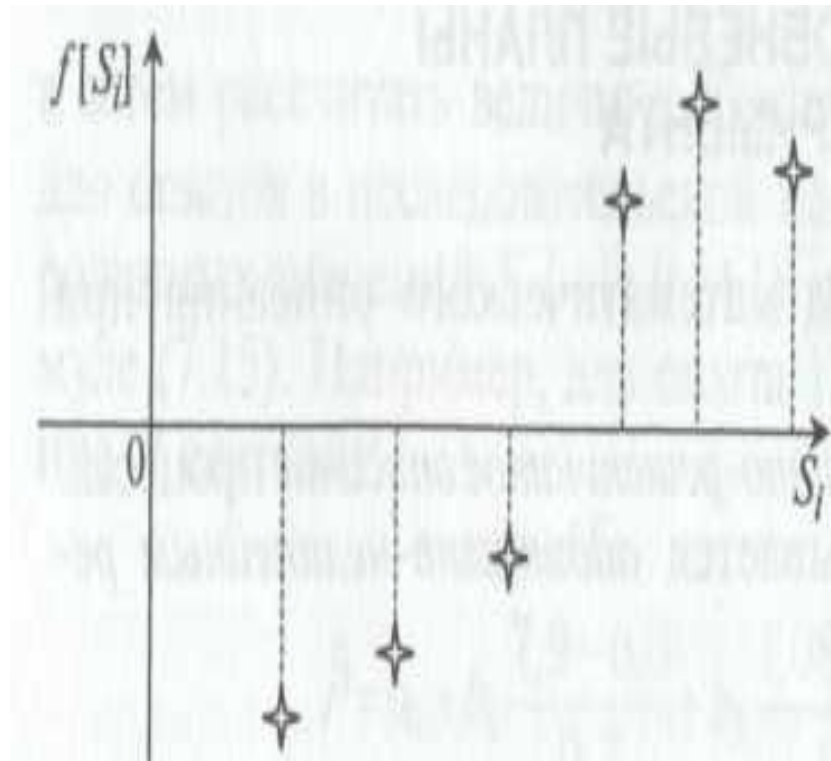
**Аддитивное** — значит состоящее из суммы членов, каждый из которых зависит только от одного из факторов.



- Принятое в методе Бокса—Уилсона линейное уравнение регрессии было аддитивным, только значения каждой функции были линейными:
- Здесь функции могут быть и нелинейными

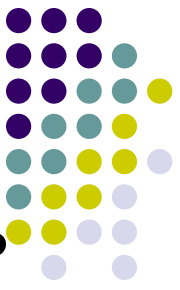
$$f_i(S_i) = b_i S_i + b_{ii} S_i^2 + b_{iii} S_i^3 + \dots$$

# Пример решетчатой функции одного фактора



$$f_i [S_i] = \begin{cases} b_{i1} & \text{при } S_i = S_{i1}; \\ b_{i2} & \text{при } S_i = S_{i2}; \\ \dots & \dots \\ b_{ik} & \text{при } S_i = S_{ik}. \end{cases}$$





- Чтобы получить зависимости с наименьшим числом опытов и упростить вычисления, целесообразно использовать *ортогональные* матрицы планирования уже на 3, 4, 5 уровнях в соответствии с выбранной разбивкой рабочего диапазона для всех факторов.
- Ортогональность матриц предполагает, что каждый уровень любого фактора сочетается одинаковое число раз со всеми уровнями остальных факторов.

На практике такие схемы планирования эксперимента называют схемами *ортогональных латинских прямоугольников* (или квадратов, если число факторов равно числу уровней каждого фактора).

Есть удобные для практики схемы планирования —  $3 \times 3$ ,  $4 \times 4$ ,  $5 \times 5$ ,  $8 \times 4$ ,  $9 \times 3$ . Первая цифра обозначает число факторов, а вторая — число уровней каждого фактора.



## Планирование эксперимента для 4 факторов на 4 уровнях

№ п/п	$S_1$	$S_2$	$S_3$	$S_4$
1	1	1	1	1
2	3	1	3	2
3	4	1	4	4
4	2	1	2	3
5	1	3	3	4
6	3	3	1	3
7	4	3	2	1
8	2	3	4	2
9	1	4	4	3
10	3	4	2	4
11	4	4	1	2
12	2	4	3	1
13	1	2	2	2
14	3	2	4	1
15	4	2	3	3
16	2	2	1	4



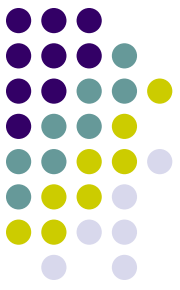
Величина эффектов решетчатого описания ( $b_{ik}$ ) и значение  $b_0$  рассчитываются почти так же просто, как в методе Бокса— Уилсона:



$$b_0 = \frac{\sum_{u=1}^N P_u}{N};$$

$$b_0 = \frac{\sum_{u=1}^N P_u}{N};$$

$P_{iu}^k$  — выход в  $u$ -м варианте плана, где  $i$ -й фактор находился на  $k$ -м уровне;  $m$  — число уровней каждого фактора



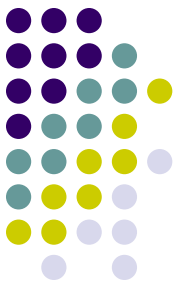
- нужно найти средний выход  $P$  в тех опытах, где вычисляемый фактор находится на определяемом уровне,
- вычесть из него средний выход по всей матрице планирования. При этом часть коэффициентов  $b_{ik}$  будет со знаком «+», а часть — со знаком «-».
- Для проверки можно использовать соотношение: для каждого фактора алгебраическая сумма  $b_{ik}$  равна 0.
- Доверительный интервал для оценки значимости коэффициентов определяется по соотношению

$$\varepsilon = t \sqrt{\frac{\sigma^2}{\gamma N / m}}$$

# Итоги расчетов величины эффектов аддитивно-решетчатого описания

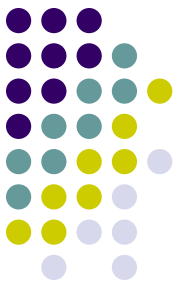


№ п/п	Наименование фактора		Натуральные и кодированные (в скобках) значения уровней факторов			
1	$S_1$ , соевая мука	Концентрация, %	1,5 (1)	2,0 (2)	3,0 (3)	4,5 (4)
		Эффект	+168	-51	-52	-67
2	$S_2$ , кукурузная мука	Концентрация, %	2,0 (1)	3,0 (2)	4,0 (3)	6,0 (4)
		Эффект	-155	+267	+92	-195
3	$S_3$ , кукуруз- ный экстракт	Концентрация, %	0 (1)	0,5 (2)	1,0 (3)	2,0 (4)
		Эффект	+62	+246	+181	-490
4	$S_4$ , мел	Концентрация, %	0,4 (1)	0,6 (2)	0,9 (3)	1,2 (4)
		Эффект	-452	+11	+108	+331



- Адекватность аддитивно-решетчатого описания проверяют аналогично методу Бокса—Уилсона. При нахождении дисперсии адекватности число степеней свободы  $f_a$  принимают  $f_a = N - p(m - 1) - 1$ .
- «Все должно быть сделано настолько просто, насколько это возможно, но не проще!»

Альберт Эйнштейн



# 1 вариант

- 1. Дайте определение понятия «биотехнология».
- 2. Какие продукты биотехнологии применяют в медицине?
- 3. Каковы сходство и различия в стадиях ферментации, биотрансформации, биокатализа?
- 4. В чем заключается классификация биотехнологических производств по типам технологических схем?
- 5. Опишите блок-схему производства лизина.
- 6. Виды продуктов по их месту в технологической схеме.