

Виды анализа и R Markdown

Способность к анализу не следует смешивать с простой изобретательностью, ибо аналитик всегда изобретателен, тогда как не всякий изобретательный человек способен к анализу.

Эдгар Аллан По, Убийство на улице Морг

Введение

Цель данной работы - привести **примеры научных работ**, в которых использовались конкретные **виды анализа** и обосновать корректность подобранных примеров. Виды анализа для моего варианта:

1. Дедуктивный (Логический)
2. Механистический

Также необходимо продемонстрировать умение пользоваться языком разметки **R Markdown**. Для этого данная работа должна быть создана с его помощью и содержать следующие элементы:

- заголовки разных уровней
- маркированные списки
- нумерованные списки
- URL-ссылки
- математические формулы в формате LaTeX
- изображения
- цитаты
- таблицы

Пойдем по порядку.

Виды анализа

Сегодня человеческая цивилизация генерирует колоссальные объемы данных, которые необходимо постоянно анализировать. Основная цель любого анализа данных – **поиск и обнаружение закономерностей** в объеме данных. Однако наборы данных бывают разные. Они отличаются объемом данных, типом и количеством показателей, полнотой, зашумленностью, описываемыми явлениями. Для разных наборов данных существуют различные виды их анализа.

В рамках курса “Методы очистки данных” были описаны 6 видов анализа данных:

1. Дескриптивный (Описательный)
2. Разведочный
3. Дедуктивный (Логический)
4. Прогностический
5. Причинно-следственный
6. Механистический

В данной работе я затрону №3 и №6.

Дедуктивный анализ

Описание

Данный вид анализа подразумевает **экстраполяцию выводов** о малой выборке на всю совокупность. Особенности дедуктивного анализа:

- Позволяет оценить значение переменной для генеральной совокупности и указывает меру неопределенности полученной оценки
- Точность выводов сильно зависит схемы выборки (выборка должна быть репрезентативна)

Пример

Мониторинг воздействия на здоровье человека производственно-технических факторов

DOI: 10.24411/1728-323X-2019-18033

В данной работе авторы исследуют влияние неблагоприятных факторов литейного производства на организм рабочих. В исследовании принимали участие 80 рабочих и контрольная группа из 50 здоровых людей. Рабочих разделили на 3 группы по продолжительности работы на производстве:

№ Группы	Стаж работы, г	Количество участников, чел	Количество участников, %
1	1 - 5	26	32,5
2	5 - 10	32	40,0
3	10 - 15	22	27,5

Рабочие были обследованы на наличие различных групп заболеваний. Большая часть рабочих (62,30%) имела заболевания органов пищеварения. Диагностика микрофлоры кишечника рабочих показала негативные отклонения от контрольной группы, выраженные в уменьшении встречаемости соответствующих бактерий (*бифидобактерий, пропониобактерий лактобактерий, бактероидов*). Важно, что отклонение увеличивалось от 1 к 3 группе. Также обнаружилось, что увеличивается встречаемость паразитов **Blastocystis spp.** от 1 к 3 группе.

Дальнейший анализ показал, что данный вид паразитов можно использовать как биологический индикатор состояния микрофлоры кишечника человека, работающего на литейном производстве.

Выводы о выборке объемом в 80 случаев были экстраполированы на совокупность всех рабочих литейных производств. Таким образом можно убедиться в использовании **дедуктивного анализа** в данном исследовании.

Механистический анализ

Описание

Выявление **точной зависимости** одних величин от других выделяет механистический анализ среди прочих видов анализа. Особенности механистического анализа:

- Данному виду анализа хорошо подходит термин “моделирование”
- Применяется к ситуациям, которые хорошо моделируются детерминированными уравнениями
- Обычно применяется к физическим и инженерным наукам (где скрытые зависимости мало влияют на объекты)
- Часто единственный шум в данных - ошибки измерения

Пример

Надежность изделий из полимерных композиционных материалов с учетом статистической изменчивости их характеристик

CyberLeninka

Данная работа посвящена изучению надежности полимерных композиционных материалов на основе полипропилена. Надежность оценивают с помощью вероятности безотказной работы R , которая рассчитывается по формуле:

$$R = 1 - \Phi \left[\frac{\bar{\sigma} - \bar{\sigma}}{\sqrt{S_{\sigma}^2 + S_{\sigma}^2}} \right]$$

, где $\bar{\sigma}$ и $\bar{\sigma}$ - среднее значение предела текучести при растяжении и среднее значение эквивалентного напряжения соответственно, S_{σ}^2 и S_{σ}^2 - среднее квадратическое отклонение предела текучести при растяжении и среднее квадратическое отклонение эквивалентного напряжения соответственно.

В данной работе оценивается влияние различных параметров на надежность (и другие величины в процессе анализа). Моделируются различные ситуации, условия работы и режимы эксплуатации. Строятся графики зависимостей одних величин от других.

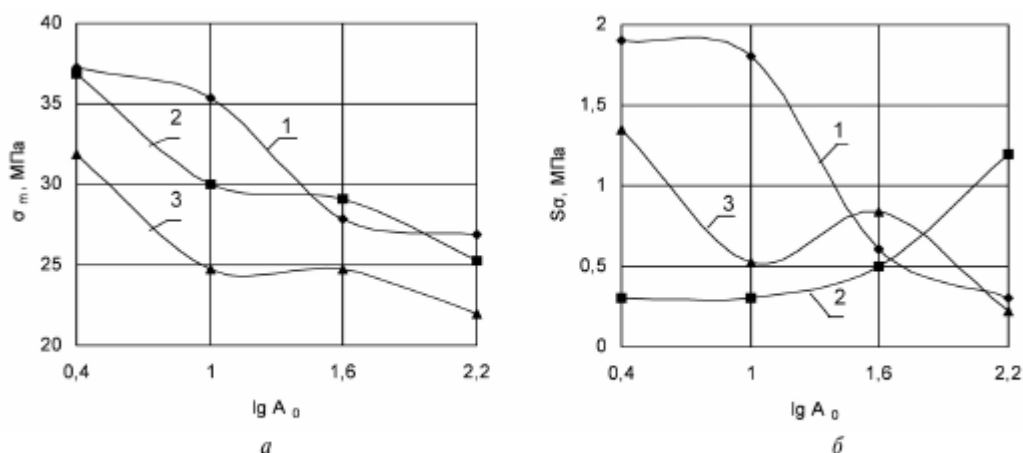


Рис. 1. Зависимость: а) среднего значения предела текучести и б) среднего квадратического отклонения предела текучести от логарифма площади поперечного сечения образцов A_0 для термопластов на основе полипропилена: 1) ПП 21060-16, бесцветный; 2) ПП 21060-A20; 3) МПП 15-04 бесцветный

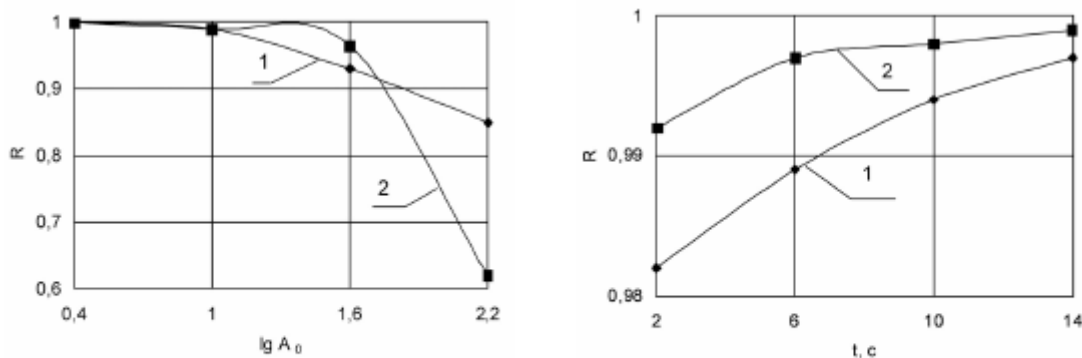


Рис. 2. Зависимость вероятности безотказной работы от \lg площади поперечного сечения образцов A_0 для термопластов на основе полипропилена: 1) ПП 21060-16, бесцветный; 2) ПП 21060-A20

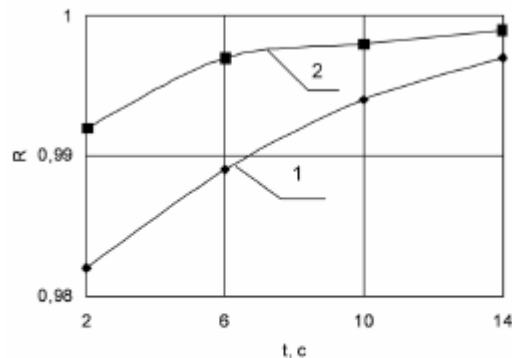


Рис. 3. Зависимость вероятности безотказной работы от времени выдержки под давлением ПП 21060-16 и температуры литья при давлении литья 80 МПа: 1) 270; 2) 200 °C

В данном исследовании выявлена точная зависимость надежности изделий из полимерных композиционных материалов от условий работы и параметров изделия, а значит использовался **механистический анализ**.

Заключение

В ходе работы были подобраны и обоснованы примеры для интересующих **видов анализа**, а также с помощью **R Markdown** был создан этот документ с внедрением необходимых структурных элементов.

Репозиторий проекта на GitHub.