## Анализ данных и стох. выч. методы (Анализ данных и стат. пакеты, 4/8, 2012/2013)

## Голяндина Н.Э., примерные темы к экзамену

- 1. 50 мини-вопросов по анализу одномерных данных
- 2. 34 мини-вопроса по анализу двумерных данных
- 3. Проверка гипотез о виде распределения. PP-plot, QQ-plot, normal probability plot. Критерии Колмогорова-Смирнова, Лиллиефорса, Шапиро-Уилка, хи-квадрат.
- 4. t-Критерий для независимых выборок. Модификации критерия. Нормальная модель и ее отсутствие. Сбалансированный дизайн.
- 5. Непараметрические критерии для сравнения независимых выборок: критерий Манна-Уитни и критерий Колмогорова-Смирнова.
- 6. Критерии для сравнения зависимых выборок: t-критерий, критерий знаков, парный критерий Вилкоксона.
- 7. Принципы сравнения критериев между собой с примерами. Сравнение критериев для проверки гипотезы о равенстве распределений.
- 8. Множественные сравнения. Поправка Бонферрони. Применение в дисперсионном анализе (роst-hoc сравнения). LSD и HSD тесты.
- 9. Множественная линейная регрессия. Проверка значимости регрессии (два вида статистики критерия на основе разложения суммы квадратов, через SSR и через множественный коэффициент корреляции)
- 10. Расстояние Махаланобиса, его распределение в нормальной модели. Проверка значимости регрессии с помощью расстояния Махаланобиса.
- 11. Оценки коэффициентов регрессии и их распределение (по исходным данным, по центрированным и по стандартизованным).
- 12. Проверка значимости коэффициентов регрессии. Влияние корреляций между независимыми признаками на точность оценивания (пример с двумя независимыми признаками). Доверительные эллипсоиды.
- 13. Теорема Гаусса-Маркова, МНК и ОМП оценки коэффициентов регрессии.
- 14. Анализ остатков в регрессионной модели, адекватность модели, hat-matrix, ортогональность остатков предсказанным значениям.
- 15. Поиск выделяющихся наблюдений в регрессии. Расстояние Кука и Махаланобиса.
- 16. Уменьшение числа зависимых переменных в регрессии. Избыточность. Пошаговая регрессия.
- 17. Интервал для предсказания и доверительный интервал для линии регрессии (среднего предсказания).