

Регрессионный анализ, часть 2. Домашнее задание, вариант 1.

Кемпбелл и Махон изучали морфологическую изменчивость (рис. 1) голубой и оранжевой форм крабов *Leptograpsus variegatus*, обитающих в Австралии (Campbell, Mahon, 1974).

Опишите зависимость максимальной ширины карапакса (CW) от толщины тела (BD) для голубой формы крабов (данные в файлах lept-o.xlsx или lept-o.csv).

- Подберите линейную модель зависимости максимальной ширины карапакса (CW) от толщины тела (BD).
- Постройте скаттерплот с линией регрессии и ее доверительным интервалом. Добавьте названия осей. Примените к графику подходящую тему оформления. Сохраните рисунок.
- Чтобы проверить условия применимости линейной регрессии, создайте датафрейм с данными для анализа остатков. Постройте и сохраните графики зависимости стандартизованных остатков от предсказанных значений и квантильный график остатков.
- Если условия применимости не выполняются, попробуйте преобразовать исходные данные и подобрать новую модель.
- Существует ли зависимость между исследованными признаками? Проверьте нулевую гипотезу о том, что коэффициент при независимой переменной равен нулю при помощи t- или F-критерия.
- Насколько хорошо подобранная линейная регрессия описывает данные? Запишите получившееся уравнение линейной регрессии и коэффициент детерминации R^2 .
- Используя подобранную модель, определите при помощи функции `predict()`, какая средняя максимальная ширина карапакса будет у крабов с шириной фронтальной области 10 и 23 мм. В каких пределах будет располагаться эта величина с 95% вероятностью? В каких пределах располагается ожидаемая ширина карапакса 95% всех крабов таких размеров?
- В отчете кратко опишите, выполняются ли условия применимости линейной регрессии и интерпретируйте результаты регрессионного анализа.
- Не забудьте проверить все ли вы сделали - посмотрите критерии оценок в главном меню на сайте. Узнать, как правильно процитировать R можно, выполнив команду `citation()`

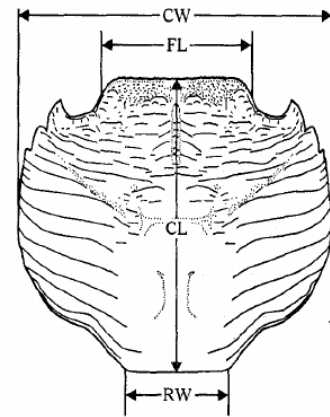


рис. 1: Карапакс краба *Leptograpsus variegatus*, вид с дорсальной стороны. Отмечены измерения **FL** - ширина фронтальной области кзади от передних фронтальных бугров, **RW** - ширина задней части, **CL** - длина по медианной оси, **CW** - максимальная ширина. Кроме отмеченных признаков, так же измеряли толщину тела **BD**. (рис. из Campbell, Mahon, 1974)