Task 2.4

Пусть t_{ij} - время, за которое нормальная температура тела восстанавливается для j-го пациента при использовании i-го одеяла ($i \in \{0, 1\}, j = \overline{1, 14}$). Формальная запись задачи:

выборки:	$t_0^{14} = (t_{01}, \dots t_{0,14}),$
	$t_1^{14} = (t_{11}, \dots t_{1,14})$
	выборки связанные
нулевая гипотеза H_0**:**	$\mu_0 = \mu_1$
альтернатива H_1**:**	$\mu_0 <\neq> \mu_1$

Один из возможных критериев для проверки этой гипотезы - t-критерий Стьюдента для связанных выборок:

статистика:	$T\left(t_0^{14}, t_1^{14}\right) = \frac{\bar{t}_0 - \bar{t}_1}{S/\sqrt{14}}$
	$S = \sqrt{\frac{1}{13}\sum\limits_{i=1}^{14}\left(D_i - \bar{D}\right)^2}$
	$D_i = t_{0i} - t_{1i}, \bar{D} = \frac{1}{14}\sum_i D_i$
нулевое распределение:	$St(13)$

Применять критерий стоит при дополнительном предположении о нормальности распределений, из которого получены данные, а также при предположении о простоте выборок. Достигаемый уровень значимости для t-критерия Стьюдента на выборке:

$$p = 2(1 - F_{St(13)}(|T|))$$

При отсутствии предположения о нормальности данных можно использовать критерий знаков для проверки нулевой гипотезы о равенстве медиан распределений для связанных выборок. Формально он строится следующим образом:

нулевая гипотеза:	$P(t_0 > t_1) = \frac{1}{2}$
альтернатива:	$P(t_0 > t_1) <\neq> \frac{1}{2}$
статистика:	$T\left(t_0^n, t_1^n\right) = \sum_{i=1}^m [t_{0i} > t_{1i}]$
нулевое распределение:	$Bin(14, \frac{1}{2})$

Для использования критерия знаков необходимо, чтобы ни для какого испытуемого данные двух экспериментов не совпали ($t_{0j} \neq t_{1j}\ \forall j$), и простота обеих выборок. Достигаемый уровень значимости:

$$p = 2(1 - F_{Bi(14, \frac{1}{2})}(|T|))$$

Недостатком этого критерия является то, что он 'выбрасывает' слишком много информации о данных в выборке. С другой стороны, он позволяет проверять гипотезу на выборках, для которых известна только нижняя или верхняя граница значения признака. Однако сложно представить, что в данной задаче могут быть цензурированные выборки.

Более мощный критерий для проверки гипотезы о равенстве медиан, не предполагающий при этом нормальности распределений - критерий знаковых рангов Уилкоксона:

нулевая гипотеза:	$\mathrm{med}\,(t_0 - t_1) = 0$
альтернатива:	$\mathrm{med}\,(t_0 - t_1) <\neq> 0$
статистика:	$W\left(t_0^{14}, t_1^{14}\right) = \sum_{i=1}^{14} \mathrm{rank}\,(\lvert t_{0i} - t_{1i}\rvert) \cdot \mathrm{sign}\,(t_{0i} - t_{1i})$
нулевое распределение:	табличное

Достигаемый уровень значимости для этого критерия будет выглядеть аналогично всем предыдущим:

$$p = 2(1 - F(|T|)),$$

однако значение $F(|T|)$ находится путем подсчета числа способов расстановки знаков '-' перед рангами при вычислении выборочной статистики, которые дают значение статистики, равное W (при условии истинности H_0 все способы равновероятны).