1. Охарактеризовать модель данных, описанных п. п. 1 –2.
2. Дана случайная выборка разделённая на две группы (сердечно-сосудистые заболевания присутствуют, отсутствуют) с атрибутами:
   1. биохимических показателей крови;
   2. артериального давления;
   3. антропометрических показателей (вес, рост, жизненная емкость легких и т. д.);
   4. факторов риска (число выкуриваемых сигарет в сутки, уровень физической активности, наличие повышенного артериального давления и т. д.);
3. Дана случайная выборка разделённая на две группы (в зоне радиоактивного загрязнения в результате аварии на Чернобыльской АЭС, на незагрязненной территории) с атрибутами:
   1. наличие различных видов заболеваний,
   2. наличие факторов риска,
   3. биохимические показатели крови, артериальное давление, иммунологические показатели.
4. Какие гипотезы можно выдвинуть, располагая этими данными?

Можно ли утверждать равенство дисперсий (мат ожиданий) для людей с сердечнососудистыми заболеваниями и без них.

Можно ли утверждать равенство дисперсий (мат ожиданий) для людей, живущих на территории радиоактивного загрязнения и не живущих там.

1. Какие допущения необходимо сделать, чтобы проверить эти гипотезы?

Данные представлены двумя независимыми случайными выборками  и  объемов  и соответственно:

= (), *i*= 1, 2. Предположим, что выборки извлечены из нормальных распределений с равными дисперсиями (= ).

1. Как установить правильность этих допущений?

Выдвинем гипотезы равенства (однородности) математических ожиданий  и дисперсий  и их двусторонние альтернативы  и 

: =, : ≠;

: = , :  ≠.

И проверим их решающим правилом, основанным на P значении: .

1. Какие критерии могут быть использованы для проверки выдвинутых гипотез?

Если необходимо выяснить, за счет чего обнаружилась неоднородность рассматриваемых выборок, то следует дополнительно произвести проверку однородности дисперсий.

**Статистический критерий однородности двух выборочных дисперсий** основан на статистике

*F*=  ,

которая в условиях справедливости гипотезы  имеет *F–*распределение Фишера с числами степеней свободы числителя и знаменателя, равными соответственно *–*1 и *–*1.

Для проверки гипотезы равенства математических ожиданий в двух выборках в этом случае используется ***t*–критерий Стьюдента**, или **двухвыборочная статистика Стьюдента**:

*t* = ( *–* ) / ( (1/+ 1/))1/2 .

Критерий оказывается чувствительным к отклонениям распределений от нормальности, различиям дисперсий  и  и неравенству объемов выборок  и .

Для обеспечения устойчивости решений к нарушению нормальности распределения и равенства дисперсий, разработаны устойчивые критерии, например**, *t*–критерий Уэлча**.

1. Как можно интерпретировать результаты проверки гипотез?

При равенстве мат ожиданий можно сказать, что:

а) сердечнососудистые заболевания не зависят от

1. биохимических показателей крови;
2. артериального давления;
3. антропометрических показателей (вес, рост, жизненная емкость легких и т. д.);
4. факторов риска (число выкуриваемых сигарет в сутки, уровень физической активности, наличие повышенного артериального давления и т. д.);

б) люди, проживающие в местах загрязнения радиоактивных веществ, имеют такой же уровень заболеваний, как и другие.

1. наличие различных видов заболеваний,
2. наличие факторов риска,
3. биохимические показатели крови, артериальное давление, иммунологические показатели.
4. Как должны быть представлены исходные данные в статистических пакетах для выполнения вычислений, связанных с проверкой гипотез согласия и однородности?

Для первого должно быть четыре колонки со значениями атрибутов, и одна колонка с двумя группами.

Для второго должно быть три колонки со значениями атрибутов, и одна колонка с двумя группами.