Статистический анализ выборок курсов валют с помощью ранговых критериев

М.А.Власов

Одна из самых обычных задач в математической статистике – это задача о сравнении двух выборок. В отличие от обычных критериев [1,8], использование ранговых критериев не накладывает никаких ограничений на тип распределения случайной величины. В ранговых критериях можно различать три группы. Критерии первой группы проверяют гипотезу о совпадении центральных тенденций сравниваемых совокупностей, второй-гипотезу о равенстве размахов варьирования, третьей- гипотезу о равенстве законов распределения[1]. Нулевая гипотеза для критериев первой группы формулируется как H_0: $\mu_\xi = \mu_\eta$, где μ_ξ и μ_η - характеристики центров распределения случайных величин ξ и η соответственно, реализациями которых являются выборочные значения x и y. Альтернативами будут: H_1: $\mu_\xi \neq \mu_\eta$; H_2: $\mu_\xi > \mu_\eta$; H_3: $\mu_\xi < \mu_\eta$.

Критерий Вилкоксона.

Пусть имеются две выборки: $x_1, x_2, \ldots x_m$ и $y_1, y_2, \ldots y_n$, причем их объёмы не обязательно должны быть одинаковыми. Объединим обе выборки в одну и упорядочим её. В качестве меры близости центральных тенденций двух выборок можно взять сумму рангов значений, принадлежащих каждой исходной выборке. Эта величина называется статистикой (критерием) Вилкоксона. Итак, имеем две величины:

$$W_x = \sum_{j=1}^{m} R(x_j^*),\ \ W_y = \sum_{j=1}^{n} R(y_j^*),\ где\ R(x_j^*), R(y_j^*) - ранги\ значений\ иксов\ и\ игреков\ в\ общей\ упорядоченной\ выборке.$$

В том случае, если объёмы выборок одинаковы, с критическим значением сравнивается меньшая сумма рангов. Нулевая гипотеза отвергается, когда W_x(или W_y)$<W(\alpha)$, где $W(\alpha)$-критическое значение статистики Вилкоксона. Для выборок с неодинаковыми объёмами вычисляют “дополнение”[1,6].

В тех случаях, когда выборка образована значениями случайной величины с нормальным распределением, лучше воспользоваться более чувствительным критерием, а именно X-критерием Ван дер Вардена[1,5].

Критерий Ван дер Вардена.

Пусть имеются две выборки: $x_1, x_2, \ldots x_m$ и $y_1, y_2, \ldots y_n$. Проделаем ту же процедуру, что и в случае критерия Вилкоксона. Разделим теперь ранги значений на $N+1$, где $N=n+m$, и вычислим сумму $X_x = \sum_{j=1}^{m} \Psi\left(\frac{R(x_j^*)}{N+1}\right)$, *где* $\Psi(\ )$ – *обратная функция стандартного нормального распределения.*

Значения X_x и X_y различаются лишь по знаку, поэтому удобнее рассчитывать статистику для выборки меньшего объёма. В зависимости от альтернативы нулевая гипотеза отвергается с уровнем значимости α, если $|X|>X(\alpha)$ для альтернатив H_2 и H_3 и если $|X|>X(2\alpha)$ для альтернативы H_1.

Критерий Манна-Уитни.

Для проверки гипотезы сдвига Манн и Уитни предложили ранговый критерий, основанный на статистике [2]:

$$U = \sum_{i=1}^{n}\sum_{j=1}^{m} h_{ij},\ \text{где}\ h_{ij} = \begin{cases} 1, & x_i < y_j; \\ 0, & x_i > y_j. \end{cases}$$

Если $U_1(\alpha) \le U \le U_2(\alpha)$, гипотеза сдвига отклоняется ($U_1(\alpha)$ и $U_2(\alpha)$- критические значения, берущиеся из таблицы).

Статистический анализ часто проводится в банковских операциях[7,9,10].

Рассмотрим примеры и применим ранговые критерии к сравнению двух выборок x и y-курсов валют за два соседних дня в девяти коммерческих банках. Были исследованы четыре базовых примера:

1) Курсы продажи доллара США за 17-18.12.2003. При этом курс Центрального Банка России(ЦБ) уменьшился, значит следует ожидать отклонение гипотезы о равенстве центральных тенденций двух выборок.
2) Курсы продажи доллара за 25-26.12.2003. В эти два дня курс ЦБ не изменился, следовательно нужно ожидать подтверждение нулевой гипотезы.

3) Курсы покупки доллара за 17-18.12.2003.

4) Курсы покупки доллара за 25-26.12.2003.(См.таблицу)

Курс доллара США к рублю.								
Банки	17.12.2003		18.12.2003		25.12.2003		26.12.2003	
	Покупка	Продажа	Покупка	Продажа	Покупка	Продажа	Покупка	Продажа
1	28	29,7	27,7	29,3	28	29,27	28	29,25
2	28,8	29,48	28,8	29,4	28,2	29,28	28,7	29,3
3	28,1	29,57	28	29,3	27,5	29,29	27	29,27
4	29	29,4	28,7	29,3	28,1	29,23	27,9	29,4
5	28,2	29,32	28,3	29,3	28	29,25	28	29,26
6	29,1	29,59	28,3	29,3	28	29,39	28	29,59
7	28,9	29,3	28,7	29,25	28,9	29,28	28,9	29,25
8	29	29,6	28,5	29,35	27,5	29,2	27,5	29,1
9	28	29,37	27,7	29,3	28	29,27	28	29,25
Курс ЦБ	29,3		29,25		29,245		29,245	

Применение к примерам 1-4 критериев дало следующие результаты:

1) Меньшая сумма рангов W=53,5<$W(\alpha)$, X=5,22> $X(\alpha)$,U=8≤ $U_1(\alpha)$ для всех α. Нулевая гипотеза отвергается. 2) W=80>$W(\alpha)$, X=0,631< $X(\alpha)$,$U_1(\alpha)\leq U=26\leq U_2(\alpha)$для всех α. Нулевая гипотеза не отвергается. 3) W=70,5>$W(\alpha)$,X=2,79< $X(\alpha)$,$U_1(\alpha)\leq U=24\leq U_2(\alpha)$для всех α. Нулевая гипотеза не отвергается.

4) W=80,5>$W(\alpha)$,X=0,844< $X(\alpha)$,$U_1(\alpha)\leq U=34\leq U_2(\alpha)$ для всех α. Нулевая гипотеза не отвергается. Критические значения брались из таблиц [1,2]. Распределения могут отличаться не только средним значением, но и разбросом, в качестве меры которого обычно выступает дисперсия. Нулевая гипотеза для критериев второй группы формулируется так:

$$H_0 : \sigma_x^2 = \sigma_y^2 \text{ } \textit{против одной из альтернатив} \text{ } H_1 : \sigma_x^2 \neq \sigma_y^2;\ H_2 : \sigma_x^2 > \sigma_y^2;\ H_3 : \sigma_x^2 < \sigma_y^2$$

в предположении, что центры распределения практически совпадают.

Модификация Сиджела-Тьюки критерия Вилкоксона

Поступим с выборками x и y так же, как и в случае критерия Вилкоксона. Ранги после этого определим по правилу чередования [4]. Обозначим S_x и S_y суммы рангов для выборок x и y соответственно. Пусть S-меньшая из сумм S_x и S_y и $S'=min(m,n)(N+1)-S$. (Здесь N=n+m). Нулевая гипотеза отвергается , если S и S' оказываются меньше критического значения $W(\alpha)$ для статистики Вилкоксона. Применение к примерам 1-4 критерия Сиджела-Тьюки дало следующие результаты:

1) S_x=75, S_y=96, S=75, S'=96> $W(\alpha)$ для всех α. Нулевая гипотеза не отвергается. 2) S_x=98, S_y=73, S=73, S'=98> $W(\alpha)$ для всех α. Нулевая гипотеза не отвергается. 3) S_x=67, S_y=104, S=67, S'=104> $W(\alpha)$ для всех α. Нулевая гипотеза не отвергается. 4) S_x=89, S_y=82, S=82, S'=89> $W(\alpha)$ для всех α. Нулевая гипотеза не отвергается.

Нулевая гипотеза для критериев третьей группы формулируется как

$H_0: F(x) = F(y)$ *против альтернатив* $H_1: F(x) \neq F(y);\ H_2: F(x) > F(y);\ H_3: F(x)$

Критерий Колмогорова-Смирнова.

Критерий основан на сравнении рядов накопленных частот обеих совокупностей. Пусть $F_j(x)$ и $F_j(y)$- накопленные относительные частоты выборок x и y; где j- номер значения в общем вариационном ряду. Максимальная по величине разность $D = \max[F_j(x) - F_j(y)]$ может служить мерой близости двух распределений. При больших объёмах совокупностей (m,n>100) если D больше критического значения $D(\alpha)$, то нулевая гипотеза о равенстве распределений отвергается. Если объёмы совокупностей малы, приходится вводить поправки [3,6]. Применение к примерам 1-4 критерия Колмогорова-Смирнова дало везде ожидаемые результаты. Таким образом можно сделать следующий вывод после применения всех ранговых критериев: лишь для примера 3 все критерии первой группы дали неверный результат. Во всех остальных случаях все критерии подтвердили ожидания.

Литература.

1. Благовещенский Ю.Н., Самсонова В.П., Дмитриев Е.А. Непараметрические методы в почвенных исследованиях. [Текст]: Монография/ Ю.Н.Благовещенский и др. - Москва: "Наука", 1987. – 98 с.
2. Кобзарь А.И. Прикладная математическая статистика. [Текст]: Монография/ А.И.Кобзарь - Москва: Физматлит, 2006, - 816 с.
3. Хеттманспергер Т. Статистические выводы,основанные на рангах.

[Текст]: Монография/ Т.Хеттманспергер - Москва:”Финансы и статистика”, 1987, - 333 с.

4. Орлов А.И. Эконометрика. [Текст]: Монография/А.И.Орлов - Москва: Экзамен, 2006, - 576 с.
5. Gibbons J. D., Chakraborti S. Nonparametric Statistical Inference, 4th Ed. — CRC, 2003 — 608 с.
6. Hajek J., Sidak Z., Sen K. P. Theory of rank tests(second edition). — Academic Press, 1999. - 450 p.
7. Пучков Е.В. Разработка системы поддержки принятия решений для управления кредитными рисками банка. [Электронный ресурс]// «Инженерный вестник Дона», 2011, №1. – Режим доступа: http://ivdon.ru/magazine/archive/n1y2011/377 (доступ свободный) – Загл. с экрана. – Яз. рус.
8. Галушка В.В., Фатхи В.А. Формирование обучающей выборки при использовании искусственных нейронных сетей в задачах поиска ошибок баз данных. [Электронный ресурс] // «Инженерный вестник Дона», 2013, №2. – Режим доступа: http://ivdon.ru/magazine/archive/n2y2013/1597 (доступ свободный) – Загл. с экрана. – Яз. рус.
9. Волосатова Т.А. Явление агрессивной скупки акций на российском финансовом рынке. [Текст] // Математические методы в современных и классических моделях экономики и естествознания: материалы регональной научно-практической конференции ППС и молодых учёных / Рост.гос.экон.ун-т “РИНХ”-Ростов н/Д., 2009. – С.142-146.
10. Данекянц А.Г. Различные модели осуществления скупки акций. [Текст] // Математические методы в современных и классических моделях экономики и естествознания: материалы региональной научно-практической конференции ППС и молодых учёных / Рост.гос.экон. ун-т “РИНХ”-Ростов н/Д., 2009. – С.146-149.