|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1. **Сходимость последовательности случайных величин**

Пусть на некотором вероятностном пространстве (Ω, F Ω, P) заданы случайные величины ξ1, ξ2, …..,ξn1,…., ξn. Величины ξ1, ξ2, …..,ξn1,…., ξn будем называть случайной последовательностью или последовательностью СВ. ξn – случайная последовательность, где n=1,2,3,…..Будем изучать сходимость последовательности ξn к случайной величине ξ при n→∞. Если последовательность ξn сходится к СВ ξ, то СВ ξ есть предел случайной последовательности ξn.Применить это определение к случайной последовательности нельзя без дополнительных предположений. Предварительно вспомним понятие сходимости неслуч. последов. S1, S2,…,Sn к S. Последов-ть случ.числа Sn сходится к пределу S, если ∀ε сущ-ет N=N(ε): |SN-S|<ε.Однако, это определение непригодно для случайной последовательности, поскольку для случайной последовательности величина |ξn - ξ| является случайной и требовать для нее выполнения условия |ξn - ξ|< ε бессмысленно без дополнительных условий.Поэтому для случайной последовательности применяются свои понятия сходимости:1. сходимость к среднему квадратичному2. сходимость по вероятности 3. сходимость с вероятностью 1 или “почти наверное ” или ”почти все”4. сходимость по распределению1. сходимость к среднему квадратичномуСлучайная последовательность ξn сходится к СВ ξ в среднем квадратичном, если выполняется условие: Е(ξn - ξ)2→0, при n→∞, где Е – мат. ожидание.Известно, что D(ξn - ξ)= Е(ξn - ξ)2-E2(ξn - ξ), то Е(ξn - ξ)2= D(ξn - ξ)+ E2(ξn - ξ), D – дисперсия.Из равенства видно, что сходимость случайной последовательности в среднем квадратичном обозначает сходимость к математическому ожиданию разности и дисперсии разности, т.е. сходимость средних значений (сходимость моментов).2. сходимость по вероятностиСлучайная последовательность ξn сходится к случайной величине ξ по вероятности, если ε>0 выполняется условие Р(|ξn - ξ|> ε)→0, при n→∞; или эквивалентное условие Р(|ξn - ξ| ≤ ε)→1 при n→∞. Этот тип сходимости означает, что при n→∞ уменьшается не только величина |ξn - ξ|, но и вероятность того, что эта величина будет малой.В терминах ε и δ это выглядет так: ε>0 существует такой номер N и такое число 0<δ<1, что Р(|ξn - ξ|> ε) < δ. Последнее условие читается так: вероятность того, что отклонение последовательности от предела > ε не превышает δ. Обозначается сходимость по вероятности так ξnξ. 3. Сходимости “почти наверняка” (по вероятности 1)Последовательность ξn сходится к ξ с вероятностью 1 “почти наверняка” если выполняется условие Р(ξn→ ξ)=1, при n→∞.Этот тип сходимости означает сходимость каждой реализации последовательности ξn  к СВ  (в каждом случае ξ может не совпадать). Для сравнения этого типа с сходимостью по вер-ти необходимо и достаточно условия сходимости: хотя бы для одного  при n→∞. для всех  при n→∞.Обозначается сх-ть как . 4. Сходимость по распределению.Случ. Последовательность ξn  сходится к ξ по распределению, если  , при n→∞, где функции непрерывны, где  - функция распределения ξn.Этот тип распределения обозначается:  ξn  ξ . Речь идет о сближении з-нов распределения. | 1. **Законы больших чисел в форме Бернулли и в форме Чебышева.**

ЗБЧ ТВ устанавливают сходимость некоторой случ. послед-ти ξn при n→∞, n и есть то большое число.Теорема Бернулли: Если p вероятность наступления события А в одном испытании Бернулли и m-число появлений события А в n независимых испытаниях Бернулли, то выполняется соотношения  при n→∞ , m\n – относительная частота события А.Теорема Бернулли утверждает, что относительная частота события А(m/n сходитсяся к вероятности появления события А (р(А)) по вероятности , т.е. , при .Теорема Чебышева: Если - последовательность независимых случайных величин с конечным мат. ожид. ,∀i, и дисперсией, ограниченной одной и той же константой ∀i, то выполняется соотношение: ,Комментарии: величина  назначается средним арифметическим СВ.- ср. арифм. Мат. Ожиданий.Теорема Чебышева утверждает, что сред. арифм. независимых случайных величин сх-ся к ср. арифм-ому их мат. ожидания по вероятности.1. **Законы больших чисел в форме Хинчина и в форме Колмогорова.**

ЗБЧ ТВ устанавливают сходимость некоторой случ. послед-ти ξn при n→∞, n и есть то большое число.Теорема Хинчина.Если  независимые случайные величины одинаково распределенные с конечным мат. ожиданием , ∀i, то выполняется соотношение ,Т.е. средн. арифм.случ.вел-н сходится к средн.арифм. их мат.ожид-ий по вероятности. Отличается она от Чебышева: здесь не требуется, чтобы случ. вВеличины имели конечную дисперсию, но надо чтобы они были одинаково распределенными.Теорема Колмогорова (Усиленный закон больших чисел):Если  независимые и одинаково распределенные случ. Величины с мат. ожид. ∀i, то выполняется соотношение:  при n→∞.Среднее арифм. случ.величин сходится к средн.арифм. их мат. Ожиданий почти наверное.1. **Порядковые статистики.**

Поря́дковые стати́стики в [математической статистике](http://ru.wikipedia.org/wiki/%D0%9C%D0%B0%D1%82%D0%B5%D0%BC%D0%B0%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B0%D1%8F_%D1%81%D1%82%D0%B0%D1%82%D0%B8%D1%81%D1%82%D0%B8%D0%BA%D0%B0) - это упорядоченная по возрастанию выборка. Это статистика, занимающая строго определенное место в ранжированной совокупности.Определение: Пусть - конечная [выборка](http://ru.wikipedia.org/wiki/%D0%92%D1%8B%D0%B1%D0%BE%D1%80%D0%BA%D0%B0) из [распределения](http://ru.wikipedia.org/wiki/%D0%A0%D0%B0%D1%81%D0%BF%D1%80%D0%B5%D0%B4%D0%B5%D0%BB%D0%B5%D0%BD%D0%B8%D0%B5_%D0%B2%D0%B5%D1%80%D0%BE%D1%8F%D1%82%D0%BD%D0%BE%D1%81%D1%82%D0%B8) , определённая на некотором [вероятностном пространстве](http://ru.wikipedia.org/wiki/%D0%92%D0%B5%D1%80%D0%BE%D1%8F%D1%82%D0%BD%D0%BE%D1%81%D1%82%D0%BD%D0%BE%D0%B5_%D0%BF%D1%80%D0%BE%D1%81%D1%82%D1%80%D0%B0%D0%BD%D1%81%D1%82%D0%B2%D0%BE) . Пусть  и . Перенумеруем [последовательность](http://ru.wikipedia.org/wiki/%D0%A7%D0%B8%D1%81%D0%BB%D0%BE%D0%B2%D0%B0%D1%8F_%D0%BF%D0%BE%D1%81%D0%BB%D0%B5%D0%B4%D0%BE%D0%B2%D0%B0%D1%82%D0%B5%D0%BB%D1%8C%D0%BD%D0%BE%D1%81%D1%82%D1%8C)  в порядке неубывания, так что .[Случайная величина](http://ru.wikipedia.org/wiki/%D0%A1%D0%BB%D1%83%D1%87%D0%B0%D0%B9%D0%BD%D0%B0%D1%8F_%D0%B2%D0%B5%D0%BB%D0%B8%D1%87%D0%B8%D0%BD%D0%B0%22%20%5Co%20%22%D0%A1%D0%BB%D1%83%D1%87%D0%B0%D0%B9%D0%BD%D0%B0%D1%8F%20%D0%B2%D0%B5%D0%BB%D0%B8%D1%87%D0%B8%D0%BD%D0%B0)   называется k-ой порядковой статистикой исходной выборки. | 1. **Генеральная совокупность. Простой случайный выбор, выборка, вариационный ряд.**

ξ - СВ, имеющая: fξ(xi) – плотность вероятности fξ(x); Fξ(xi) – ф-ция распределения Fξ(x).Вектор: ξ=(ξ1,…,ξn).Эксперимент над случ. величиной ξ с функцией распределения fξ(x) состоит в том, что мы осуществляем комплекс условий в результате которого случайная величина принимает свое значение и фикс-ем это значение x. Выполнив это n раз, получим n чисел: x1…xn. Значение хi, i=1бт называется *выборочным* или наблюдаемым значением ξ. Все множество возможных значений ξ наз-ся генеральной совокупностью, а множ-во чисел x1…xn – *выборкой из ген-ой сов-сти* Fξ(x).Если эксперимент организован т. о. что вер-ть быть выбранным для каждого значения xi, одна и та же и эксперименты друг от друга независимы, то такой выбор называется *простым случайным*. Для прост. Случ. выб-ки x1…xn независимы и р(xi)=1/n.Мат.статистика занимается разработкой методов получения научно-обоснованных выводов о СВ по результатам экспериментов, т.е. по выборке x1… x n. Для этого выполняется обработка выборки. Простейшая состоит в ее сортировке (х1<x2<…<xn). Ряд x1… x n называется *вариационным рядом* или *простым статистическим рядом.* хi – i-я *порядковая статистика*; х1 – минимальное выборочное зн-ние (1-я порядковая статистика).Подходы к выборке (взгляды на выборку). Статистика.Существуют два взгляда на выборку x1…xn и распределение Fξ(xi) (fξ(xi)) (пред-тся простой случ.выбор).1-й взглядВыборочные значения рассматриваются как возможные значения некоторой СВ ξ\*, имеющие одинаковые вероятности р(хi)=1/n. Ряд распределения такой СВ имеет вид:

|  |  |  |  |
| --- | --- | --- | --- |
| Xi | X1 | … | Xn |
| pi | P1 | … | Pn |

Ф-ция распр-ния ξ\* наз-ся *эмпирической* или *выборочной* ф.распр-ния СВ ξ и обозн-ся F\*ξ(x), т.о. F\*ξ(x) = Fξ\*(x)= Р(ξ\*<x). Т.о. ф.распр-ния F\*ξ(x)=m/n, где m – число выборочных зн-ний, меньших х.Используя порядковые статистики, эмпирич-кую ф-цию можно записать след.обр.: F\*ξ(x) = 0 при х<x(1), F\*ξ(x) = (i-1)/n при x(i-1)<=x<x(i), F\*ξ(x) = 1 при x>=x(n).2-й взглядиз следующих рассуждений: в результате n экспериментов мы получаем n чисел x1,…,xn или точку в n-мерном пространстве. Ничто не мешает выполнить еще n экспериментов => новая точка в n-мерном пространстве, отличная от первой. Такие эксперименты можно повторять много раз и получить много точек в n-мерном пространстве. Каждую точку рассмотрим как реализацию случайного вектора (x1,…,xn). Т.о. выборку можно рассматривать как случайный вектор, т.к. выбор простой случайный, то можем опр. закон распределения этого случайного вектора по теореме умножения, для независимых случайных велечин: или , где fξ(xi) – плотность вероятности fξ(x)Fξ(xi) – ф-ция распределения Fξ(x)Если считать выборку как случ. вектор, то любая статистика g(x1,…,xn) будет случайной величиной как ф-ция случайного вектора.В силу 2-го подхода к выборке статистикой явл-ся ф-ция случ.вектора и сама явл=ся СВ-ной. Поэтому можно говорить о ее ф-ции распределения: Fg(x). Или о ее числовых характеристиках.ПРИМЕР:1. Z=1/n\*∑xi – среднее арифм.выбор.значений – выборочное среднее (так наз-ся в МС).2. Мат.ожидание любой статистики определяется:  |
| **5. Точечные оценки характеристик и параметров распределений, их свойства.***Характеристики распределения* – ф-я распределения, плотность вер-ти, моменты и др.*Параметры распределения* – те величины, от которых зависти ф-я распределения или плотность вер-ти.Нормальное распределения с плотностью вероятности  имеет 2 параметра: a, δ2, которые являются мат. ожиданием и дисперсией. Такое распр. обазнач. N(a, δ2).δ2 = D(ξ) – параметр масштаба. а = Е(ξ) – параметр сдвига. Экспоненциальное распределение: пл-ть вер-ти , λ - параметр, λ = Е(ξ).Обозначения распределений: N(a, δ2), Е(λ), U(a,b)-равномерное.Одной из задач мат. статистики является получение по выборке х1…хn из распределения Fξ(x) оценок характеристик или параметров распределения.Оценки обозначаются такими же буквами что и параметры но с крышечкой.ПРИМЕР: - оценка параметра а нормального распределения . N(a, δ2). Оценки должны обладать свойствами близости (быть ближе) к оцениваемым параметрам. Качество оценок хар-ся такими свойствами как *несмещенность, состоятельность, эффективность*. Хорошая оценка должна обладать ими или частью их.1. Оценка  является несмещенной, если ее мат. ожидание совпадает с параметром, т.е. . Несмещенная оценка - это оценка точная в среднем, т.е. не имеющая систематической ошибки.Величина  называется смещением оценки, смещение зависит от параметра. Если мы имеем оценку, то предполагаем, что параметр неизвестен, в этом случае смещение явл-ся функцией параметра. Оценка несмещенная, если смещение = 0.Оценка является асимптотически несмещенной если b(Q)→0 при n→∞. Если оценка имеет линейное смещение, т.е. b(Q)=ą + β\*Q, то такое смещение можно устранить. Новая оценка: =(-ą)/(β+1) будет несмещенной. Оценка  называется исправленной.2. Оценка  параметра Q называется состоятельной, если она сходится к параметру Q по вер-ти, т.е. если →Q при n→∞ по вероятности, или в развернутом виде: P(|-Q|>ε)→0 при n→∞.Состоятельная оценка – оценка, точность которой улучшается при увеличении объема выборки.Теорема: несмещенная оценка  параметра Q является состоятельной, если  при n→∞.Оценка  называется строго состоятельной, если она сходится к Q, почти наверное P(-Q, при n→∞)=1.3. Оценка  называется эффективной, если она имеет min *вариацию* V()=E((-Q)2) по сравнению с любыми др. оценками. Величина *l*=(Е(эф-Q)^2)/(Е(-Q)^2) называется эффективностью оценки. В числителе – вариации эффективности. Ясно, что 0≤*l*≤1. эф – эффективная неизвестная нам оценка.Если *l*→1, при n→∞, то оценка называется асимптотически эффективной.Если несмещенность и состоятельность встречаются часто, то эффективность – значительно реже. | **6.** **Неравенство Рао-Крамера.**В [математической статистике](http://ru.wikipedia.org/wiki/%D0%9C%D0%B0%D1%82%D0%B5%D0%BC%D0%B0%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B0%D1%8F_%D1%81%D1%82%D0%B0%D1%82%D0%B8%D1%81%D1%82%D0%B8%D0%BA%D0%B0) неравенством Краме́ра — Ра́о (в честь [Гаральда Крамера](http://ru.wikipedia.org/w/index.php?title=%D0%9A%D1%80%D0%B0%D0%BC%D0%B5%D1%80,_%D0%93%D0%B0%D1%80%D0%B0%D0%BB%D1%8C%D0%B4&action=edit&redlink=1) и [К. Р. Рао](http://ru.wikipedia.org/wiki/%D0%A0%D0%B0%D0%BE%2C_%D0%9A%D0%B0%D0%BB%D1%8C%D1%8F%D0%BC%D0%BF%D1%83%D0%B4%D0%B8_%D0%A0%D0%B0%D0%B4%D1%85%D0%B0%D0%BA%D1%80%D0%B8%D1%88%D0%BD%D0%B0)) называется неравенство, которое при некоторых условиях на [статистическую модель](http://ru.wikipedia.org/w/index.php?title=%D0%A1%D1%82%D0%B0%D1%82%D0%B8%D1%81%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B0%D1%8F_%D0%BC%D0%BE%D0%B4%D0%B5%D0%BB%D1%8C&action=edit&redlink=1) даёт нижнюю границу для [дисперсии](http://ru.wikipedia.org/wiki/%D0%94%D0%B8%D1%81%D0%BF%D0%B5%D1%80%D1%81%D0%B8%D1%8F_%D1%81%D0%BB%D1%83%D1%87%D0%B0%D0%B9%D0%BD%D0%BE%D0%B9_%D0%B2%D0%B5%D0%BB%D0%B8%D1%87%D0%B8%D0%BD%D1%8B) [оценки](http://ru.wikipedia.org/wiki/%D0%A2%D0%BE%D1%87%D0%B5%D1%87%D0%BD%D0%B0%D1%8F_%D0%BE%D1%86%D0%B5%D0%BD%D0%BA%D0%B0)неизвестного параметра, выражая её через [информацию Фишера](http://ru.wikipedia.org/wiki/%D0%98%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%86%D0%B8%D1%8F_%D0%A4%D0%B8%D1%88%D0%B5%D1%80%D0%B0). Формулировка: Пусть дана статистическая модель – выборка размера n, определена функция правдоподобия и выполнены следующие условия регулярности:1. и везде [дифференцируема](http://ru.wikipedia.org/wiki/%D0%94%D0%B8%D1%84%D1%84%D0%B5%D1%80%D0%B5%D0%BD%D1%86%D0%B8%D1%80%D1%83%D0%B5%D0%BC%D0%B0%D1%8F_%D1%84%D1%83%D0%BD%D0%BA%D1%86%D0%B8%D1%8F) по .
2. Функция   ([функция вклада выборки](http://ru.wikipedia.org/w/index.php?title=%D0%A4%D1%83%D0%BD%D0%BA%D1%86%D0%B8%D1%8F_%D0%B2%D0%BA%D0%BB%D0%B0%D0%B4%D0%B0_%D0%B2%D1%8B%D0%B1%D0%BE%D1%80%D0%BA%D0%B8&action=edit&redlink=1)) имеет конечную [дисперсию](http://ru.wikipedia.org/wiki/%D0%94%D0%B8%D1%81%D0%BF%D0%B5%D1%80%D1%81%D0%B8%D1%8F_%D1%81%D0%BB%D1%83%D1%87%D0%B0%D0%B9%D0%BD%D0%BE%D0%B9_%D0%B2%D0%B5%D0%BB%D0%B8%D1%87%D0%B8%D0%BD%D1%8B) (или, что то же, конечна [информация Фишера](http://ru.wikipedia.org/wiki/%D0%98%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%86%D0%B8%D1%8F_%D0%A4%D0%B8%D1%88%D0%B5%D1%80%D0%B0)).
3. Для любой статистики  с конечным вторым моментом имеет место равенство .

Пусть при этих условиях дана статистика , которая несмещённо оценивает дифференцируемую функцию . Тогда справедливо следующее неравенство: , где ; равенство достигается тогда и только тогда, когда .Здесь  — количество [информации по Фишеру](http://ru.wikipedia.org/wiki/%D0%98%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%86%D0%B8%D1%8F_%D0%A4%D0%B8%D1%88%D0%B5%D1%80%D0%B0)  в одном наблюдении, а   - плотность распределения генеральной совокупности X в случае непрерывной статистической модели и вероятность события (X = t) в случае дискретной статистической модели.Применение: оценка параметра называется [эффективной](http://ru.wikipedia.org/wiki/%D0%AD%D1%84%D1%84%D0%B5%D0%BA%D1%82%D0%B8%D0%B2%D0%BD%D0%B0%D1%8F_%D0%BE%D1%86%D0%B5%D0%BD%D0%BA%D0%B0), если для неё неравенство Крамера — Рао обращается в равенство. Таким образом, неравенство может быть использовано для доказательства того, что дисперсия данной оценки наименьшая из возможных, то есть что данная оценка в некотором смысле лучше всех остальных.**19.Распределение Фишера**,  - независимые случайные величины. ,  имеют распределение Фишера с m, n степенями свободы и обозн-ся . Плотность вероятности распределения Фишера имеет вид:это нессиметричные кривые, распол.на положитительной полуоси абсцисс и достигающие max вблизи точки x=1. Для практики разработаны таблицы процентных отклонений этого распределения. . Распределение Фишера имеет след. св-во: Если ,то . | **7,8. Выборочные числовые характеристики одномерного и многомерного распределений.**Пусть имеем выборку x1…xn из распределенпия fξ(x), имеющего мат. ожидании а, дисперсию δ2, начальные моменты v1,v2…. и центральные моменты μ1, μ2,… Эти параметры считаются независимыми. Требуется получить оценки числовых характеристик этого распределения. Выборочные числовые характеристики явл-ся оценками соответствующих перечисленных теоретических характеристик.Выборочный метод получения оценки состоит в том, что вместо теоретической (генеральной) рассматриваем эмпирическую . Эмпирическая функция распределения является оценкой теоретической.Числовые характеристики эмпирической ф-ции распределения явл-ся оценками соответствующих хар-к генеральной ф-ции распределения. Полученные таким образом оценки называються выборочными и обозначаються теми же буквами, что и генеральные, но с чертой сверху. Выборочные нач. моменты определяються выражением: - нач.момент k-го порядка;  , =. Обозначим =-выборочное среднее. Выборочные центральные моменты:,k=1,2,…; . наиболее важными явл-ся выб.центр.мом.2-го порядка: - выборочная дисперсия. Рассмотрим метод, когда выборочный метод извлекается из 2-мерного распределения fξ(x,у). В этом случае выборка представляет сабой n пар чисел: (),…,(), где -наблюдение первой СВ -наблюдения второй СВ . Теперь можем определять выборочное среднее каждой компоненты: ; . Выборочная дисперсия: , . Выборочный коэффициент ковариации: . Выборочный коэффициент корреляции: . Привиденные выборочные числовые характеристики являються оценками соответствующих теоретических хар-к. Св-ва выборочнх числовых хар-к**10.Эмпирическая ф-ция распределения**Эмпирическая (выборочная) ф-ция распределения является оценкой теоретической ф-ции распределения по выборке x1,…, θn.Эмпирической ф-цией распределения F\*(x) называется ф-ция распределения дискретной случайной величины с возможными значениями равными выборочным значениям x1,…, θn и их вероятностями 1/n. Такую величину обозначим как ξ\*.По определению F\*(x)=F(x)Эмпирическая ф-ция распределения определяется след. выражением: F\*(x)=m/n, где m – число выборочных значений меньше x, n – объем выборки. |
| **11.Гистограмма**Гистограмма – это оценка теоритич. (генеральной) пл-ти вер-ти . Гистограмма строится так:1) весь диапазон x(1)- x(n) делится на L интервалов (L=10-12) с длинной ∆i каждая i=1,L. Подсчитывается число выборочных значений mi, попавших в i-тый интервал. ∑mi=nНа каждом интервале, как на основании, строится прямоугольник с высотой , - длина i-го интнрвала, n- объем выборки.**теорема:** пусть  в точке разбиения в обл.возможных зн-ний,  кол-во выборочн. зн-ний, попавших в i-ый интервал, n- объем выборки. Если максимальная длина интервала ->0 при n -> бесконечности, то гистограмма  явл-ся состоятельной оценкой плотности вер-ти, т.е.для  и  выполняется соотношение:Если длины интервалов не уменьшаються, то гистограмма не будет состоятельной оценкой. **18.Распределение Стьюдента.**Случайная величина ,  - независимые случ. величины и , , имеет распределение, которое наз-ся распределением Стьюдента с n степенями свободы и обозначается .Функция пл-ти вер-ти этого распределения имеет вид:, . Кривые плотности вер-ти изображены на рис: При n>30  близка к пл-ти вер-ти ; при малых n заметно от его отличается; при n=1 это неизвестное распределение Коши, которое известно тем, что не имеет моментов; при n>2 моменты существуют и равны .Для практических приложений  таблицы процентных отклонений этого распределения. Эти таблицы позволяют решить уравнение . | **12.Методы получения точечных оценок параметров распределения**Постановка задачи: будем рассм-ть скалярные случ. величины.пусть имеется выборка x1,…,xn из некоторого распределения, известного сточностью до = Векторного параметра: Требуется найти оценку  этого векторного параметра. Т.е. будет векторной случайной величиной .Для решения этой задачи существуют методы:1. метод моментов
2. максимума правдоподобия
3. максимума апостериорной плотности вер-ти
4. байесовский метод

Метод моментовМетод моментов основан на исп-нии выборочного метода. Состоит в след.:а) находятся m теоретических (генеральных) моментов , . Эти моменты явл-ся функциями неизвестных пар-ров, т.к. плотность вер-ти генеральной совокупности  известна с точностью до пар-ра;б) находим m выборочных начальных моментов , ; Приравниваем к теоретическим и выборочным моментам, в рез-те получаем систему уравнений: ,  относительно параметров . Решение этой системы даст нам оценки  по методу моментов.Св-ва получ. оценок иссл-ют отдельно. Метод рекомендуется для получения не более 2-3-х оценок. Достоинства: простота.Недостатки: не доказано, что полученные оценки обладают хорошими св-вами. Вопрос св-в полученных оценок решается отдельно.**32. Критерий согласия λ (Колмогорова)**Пусть Fξ(x) – функция распределения генер. совокупности. F0(x) – гипотетическая функция распределения, она должна быть непрерывной. Проверяется Используется статистика: , где  - максимальное по модулю отклонение гипотетической функции распределения F0 от эмпирической функции распределения F\*(x). 1Если Ho верно, то λ имеет распределение, приближающееся при n→∞ к некоторому ассиметричному распределению – Колмогорова. Критерий – правосторонний вида: λą определяется по таблицам распределения Колмогорова: ą=0,01 : λ=1,627 и т.д. | **14.Метод максимума апостериорной плотности вероятности.**в методе считается, что известно с точностью плотность вероятности . Параметр  –случайный вектор, который принял на время извлечения выборки какое-то значение, которое и требуется оценить.Известна плотность вероятности вектора  это Требуется по выборке  найти оценку Эта задача относится к классу байесовских задач. Она характеризуется тем, что считается известными распределения неизвестных величин.Метод состоит в том, что оценка отыскивается из условия максимальной апостериорной плотности вероятности:Можно максимизировать логарифмич. апостериорную плотность вероятности: Апостериорная плотность вероятности определяется по формуле Байеса: , Т.к. знаменатель в формуле Байеса не зависит от параметра , то он не влияет на результат максимизации апостериорной плотности вероятности достаточно решать задачу  или **15.Байесовский метод**Как и в методе максимума апостариорной плотноси вероятности считается известна с точностью плотность вероятности . Параметр –случайный вектор, который принял на время извлечения выборки какое-то значение, которое и требуется оценить.Известна плотность вероятности вектора это . Требуется по выборке  найти оценку , минимизирующую средний риск Средним риском наз-ся мат. ожидание функций потерь. Байесовский метод состоит в том, что оценки опред-ся из условия минимума среднего риска,т.е. как решение На практике применяется квадратичная функция потерь. Байесовская задача: минимизация условного риска R, где -апостериорная плотность вероятности параметра , определяется по формуле байеса, т.е. алгоритм: 1) записать функцию правдоподобия . 2)найти . 3)найти условный риск R . 4) найти  из условия R→min.Решение задачи упрощается при квадратичной функции потерь, при оценивании скалярного параметра:Найти: параметра , при Задача сводиться к минимизации условного риска. |
| **17.Распределение χ2**В мат. статистике широкое применение находят распределения , Фишера и Стьюдента. распределение вида , где  независимая нормально распределенная с 0-вым средним и единичной случайной величиной, распред. по одному и тому же закону *N(0,1)* наз-ся распределением с n степенями свободы, а распределение этой величины наз-ся распределением с n степенями свободы .Приведем плотность вер-ти: , где  - гамма ф-ция.Графики плотности вероятности этого распределения:  ,  , .Эти несимметричные кривые, расположенные на положит.полуоси абсцисс. Любая кривая имеет 1максимум в точке x=k-2.Существуют таблицы этого распределения, они позволяют решить ур-ние .Величина наз-ся  процентным отклонением.  - 5% отклонение. Очевидно следующее свойство: если  имеет распределение  и  и они независимы, то .**28. Проверка гипотез о дисперсии нормальной генеральной совокупности**при известном математическом ожидании:По выборке х1, … ,хn, N(a,δ2) проверить Н0: δ2=. - гипотетическое значение дисперсии, – известно. Для проверки гипотезы используется статистика: , если Н0 верна, т.е., . Н1: ; ;Н1: ; ; Н1: ; ; ; Проверка гипотезы при дисперсии N распределении при неизвестном математическом ожидании.х1, … ,хn, N(a,δ2) Н0: , а – неизвестна.Используется статистика . Если Н0 верно, то  Гипотеза проверяется аналогично предыдущему случаю и распределения H1(n) на H1(n-1).  | **20. Распределение выборочного среднего и выборочной дисперсии нормальной генеральной совокупности.**Пусть  выборка из распределения . Известны следующие статистики:  - оценка мат. ожидания;  - выборочная дисперсия;  - несмещенная оценка;  - выборочная дисперсия при известном мат. ожидании а, то имеет место след. распределение статистик:1. 2. , 3. 4. 5. При этом вел-ны V и U – независимы.1. Нормальность распределения  следует из того, что это лин. ф-ция норм.СВ xi. А лин.ф-ция не изменяет весь закон распределения.2. U – приведение  к центральному (-а) и нормированному виду (корень(n)/δ)3. z = ∑Ui^2, где Ui = (xi-a)/δ => Ui принадлежит  => по определению z принадлежит .4. Если для получения некот. статистики вместо неизвестных параметров используются их оценки, то число степеней своюоды уменьшается. Поскольку в выражении для V вместо а исп-ся , то число степеней свободы уменьшается на 1.5. t – ничто иное как определение распределения Стьюдента: **33. Критерий согласия ω2 (Мизеса-Смирнова)**, F0(x) – известная гипотетическая функция распределения.В качестве меры отклонения гипотетических данных от эмпирических используется величина ω2: (1) - это взвешенное интегральное квадратичное отклонение гипотетической функции распределения F0(x) от эмпирической F\*(x). Xk – k-я порядковая статистика. Для проверки гипотезы используется статистика z=nω2:Критерий для проверки гипотезы nω2 – правосторонний, где P(z>zα)=α. ą=0.01 и z = 0.74 и т.д. | **21. Интервальные оценки параметров распределения, методика построения доверительного интервала**Интервальной оценкой некоторого параметра наз-ся интервал  накрывающий параметр  с некоторой вер-тью , т.е. интервал удовлетворяющий ур-ю: (1).  наз-ся доверительной вероятностью или коэффициентом доверия, она выбирается близкой к 1 из набора величин 0.9, 0.95, 0.975.Величины  наз-ся нижней и верхней границами доверительного интервала или доверительными пределами. Величина (*l*=Qв-Qн) длинна доверительного интервала. Параметр величина не случайная, а интервалы случайные.Задача построения доверительного интервала для параметра формируется следующим образом: известна плотность вероятности генеральной совокупности с точностью до параметра , т.е  известна и выборка x1,…,xn из этой совокупности найти: интервал, удовлетворяющий (1);  выбираем сами.Решение: решить задачу значит решить ур-е (1) с двумя неизвестными  существует бесконечное множество доверительных интервалов. Для придания задаче однозначности от ур-я (1) переходят к двум ур-ям вида:  (2). Смысл (1) и (2) иллюстрируется рисунком: Причем: . Обычно строим симметричный доверительный интервал, когда вер-ти в системе (2) равны между собой: .**27.Проверка гипотез о математическом ожидании нормальной генеральной совокупности**при известной дисперсии:Известно: х1, … ,хn, N(a,δ2), δ2 – известнапроверить: Н0: а=а0. Для проверки этой гипотезы ипользуется статистика:P(|U|>Uα/2)= ; H1: аа0 – двусторонний критерийКритерий значимости правосторонний имеет вид: р(U>Uα)=α. H1: а>a0Левосторонний имеет вид: : р(U<Uα)=α. H1: а<a0Эмпирическое значение статистики рассчитывается по формуле: . Проверка гипотез о мат. ожидании нормальной генеральной совокупности при неизвестной дисперсии:х1, … ,хn, N(a,δ2) δ2 – неизвестна, Н0: а=а0 применяется статистика: H1: а>a0 ; P(t>tα)= правосторонняя;H1: а<a0 ; P(t<tα)= α левосторонняя;H1: аa0 ; P(|t|>tα/2)= α двусторонний;  отклонение распределения. |
| **22. Доверительный интервал для мат. ожидания нормальной генеральной совокупности.**При известной дисперсии:Для построения этого интервала используется статистика . В соответствии с методикой построения доверительного интервала решаем 2 уравнения:  Обычно Uн, Uв обозначаются индексом (1-γ)/2. Величина (1-γ)/2 определяется с помощью таблиц процентных отклонений распределения N(0,1). Но поскольку для нормального распределения нет таких таблиц, необходимо воспользоваться другой таблицей.и получим Р(-<<) = γ или Р(-<<) = γ.Отсюда получаем доверительный интервал: . Его можно записать в виде . , где (1-γ)/2 =  - процентное отклонение распределения . Длинна интервала: *l* = .При неизвестной дисперсии:х1…хn, . Дисперсия (сигма в квадрате) неизвестна, требуется построить доверительный интервал для параметра а. Получаем статистику U как и раньше, но теперь она зависит от неизвестного параметра  и не может исп-ся для построения доверит-го интервала. Поэтому предложим поиск подходящей статистики. - эта статистика не содержит  и известен ее закон распр-ния.Применяя изложенную методику (т.е. просто заменяем ина и)получаем :, где  - это отклонение . | **23. Доверительный интервал для дисперсии нормальной генеральной совокупности**При известном мат. ожиданииПусть имеется выборка из , а известно. Найти интервальную оценку дисперсии.Для построения интервала используется статистика: .Эта статистика не содержит др неизвестных параметров, кроме , известен ее закон распределения (имеет распределение ), поэтому она может использоваться для построения доверит. интервала.Uв находится непосредственно из таблиц процентных отклонений. Т.к. таблицы процента отклонений распределения  позволяют решить только первое из этих уравнений, то 2-ое ур-ние необходимо привести к виду первого, т.е. решать ур-ия: ; Р(<<) = γ ;  По изложенной методике получили интервал: При неизвестном мат. ожидании.Пусть имеется выборка из , а неизвестно. Найти интервальную оценку дисперсии.Для построения интервала используется статистика: . Аналогично предыдущему получаем интервал:  | **24. Доверительный интервал для вероятности появления случайного события**Рассмотрим случай события А; р(А)=р. Выполним n экспериментов (независимых испытаний Бернулли). Пусть в результате событие А появилось *m* раз. Известно, что  называется относительной частотой испытаний и является точной оценкой вероятности *р.* Наша задача состоит в построении интервальной оценки. Будем исходить из того, что n – большое.Построим доверительный интервал вида  с вероятностью . Для этого надо иметь статистику.На основании интегральной предельной теоремы Маура-Лапласа можно записать, что: при большом *n*  имеет распределение N(p, ), q=1-р.Возьмем  - это известно. Можно использовать для построения доверительного интервала. Применяя описанную методике можно получить доверительный интервал:Где =1-m/n.( - это \*100% отклонения распределения N(0,1), т.е.реш-ния ур-ния ).**25.Определение статистической гипотезы. Классификации гипотез.**статистической гипотезой называется непротиворечивое множество утверждений{} относ-ся распределения генеральной совокупности. Такая гипотеза называется *k-*альтернативной.Гипотеза вывода {} наз-ся двухальтернативной. Каждое утверждение многоальтернативной гипотезы наз-ся альтернативой гипотезы или гипотезой.Проверить гипотезу –это значит по выборке () сделать научнообоснованное заключение о том, какая из из альтернатив является истинной. Если какая-то из альтернатив принята истинной, то остальные считаются ложными. Обычно проверяется истинность нулевой гипотезы, т.е. альтернативы .Если в результате проверки гипотезы альтернатива принята истинной, то это не означает, что доказана математически истинность этой альтернативы.Проверка гипотезы выполняется с помощью критериев проверки гипотезы. Критерий проверки – это правило, с помощью которого можно установить истинность той или иной альтернативы.Альтернатива: простые и сложные.Альтернатива называется простой, если она однозначно определяет распределение генеральной совокупности и сложной в противном случае. Гипотеза многоальтернативная простая, если все простые, и сложной, если хотя бы одна альтернатива сложная.Гипотезы: параметрические и непараметрические. наз-ся параметрической, если она определяется значением параметра. Если утверждение относится к закону распределения в уелом, а не кего параметрам. То эта альтернатива наз-ся непараметрической. Если все непараметрические, то она непараметрическая.Проверяется гипотеза против альтернативы , т.е.. . |
| **26.Критерий значимости для проверки двухальтернативной сложной гипотезы**Проверка – двухальтернативная сложная гипотеза, Н­0 – простая, Н1 – сложная. Такая гепотеза проверяется с помощью так называемого критерия значимости. В основе критерия лежит некоторая статистика, представляющая собой отклонение гипотетических данных от эмпирических. Поскольку выборочные значения являются случайными величинами, то при любой выборке мы будем иметь какое-то значение статистики g = g(х1...хn ). Поскольку g – отклонение теоретических данных от эмпирических, то наша задача состоит в определении того, значимо ли для нее полученное значение статистики или незначимо и им можно принебречь. Пусть известна плотность вероятности статистики g. Она должна быть известна при условии, что верна. Критерий значимости записывается так: (1);или в виде *P(g> )=* (2) или *P(g<)=* (3). наз-ся уровнем значимости, выбирается статистиком из набора малых чисел {0.1,0.05,0.025 }. При таком выборе событие в формулах (1),(2),(3) является практически невозможным. Величины и наз-ся пределами значимости. Области, определённые неравенствами в формулах (1),(2),(3) наз-ся критическими областями. Области отмечены на рисунках:Критерий значимости (1) наз-ся критерием с двухсторонней критической областью, или двухсторонним критерием. Критерий (2) – правосторонний, (3) –левосторонний. Выбор того или иного критерия зависит от альтернативной гипотезы H1. Например, если для параметрической гипотезы относительно параметра θ: , то применяется двухсторонний критерий; если , применяется правосторонний критерий; если - левосторонний критерий.Гипотеза с его помощью проверяется так:1). Выбираем α=0,05; 2). По таблице процентных отклонений распределения статистики g находим предел значимости gα/2 или gα ; 3). По функции для статистики g находим эмпирическое значение статистики gэ ; 4). Сравниваем gэ и gα. Если - случай 2-хстороннего критерия (1). Если gэ>gα , в случае критерия (2), если gэ<gα , в случае критерия (3), то проверяемая гипотеза Н0 отклоняется в пользу альтернативы H1. Если эмпирическое значение статистики попадает в критическую область, то проверяемая гипотеза Н0 отклоняется. Гипотеза в этом случае отклоняется потому, что произошло событие, которого не должно быть. В этом случае есть противоречие между эмпирическими и гипотетическими данными. Величина является вероятностью принятия верной гипотезы и называется мощностью критерия. Методы, которые для каждой выборки формально точно определяются, удовлетворяют выборочные данные нулевой гипотезы или нет, называются критериями значимости. | **29.Проверка гипотезы о равенстве дисперсий двух нормальных генеральных совокупностей с неизвестным математическим ожиданием**Имеются две выборки х1, … ,хm, N(a1, ) и y1, … ,yn из распределения N(a2, ). Параметры распределения неизвестны. Требуется проверить гипотезу , когда а1, а2 - неизвестны. Для проверки гипотезы используется статистика вида . Если проверяемая гипотеза Н0 верна, то f имеет распределение F(m-1,n-1) (Фишера)(не содержит неизвестных параметров и ирименяется для проверки гипотез).При альтернативе используется правосторонний критерий значимости .Иногда для проверки подобных гипотез используют не выборочные гипотезы, а так наз. суммы квадратов , . Понятно, что , . В этом случае статистика имеет вид: .**31. Критерий χ2 (Пирсона).**Задача: по выборке х1, … ,хn из генеральной совокупности  требуется проверить гипотезу  , где  - гипотетическая плотность вероятности, известная с точностью до m параметров θ1, … , θm.Если гипотетическая плотность вероятности известна полностью, то считаем m=0. Это α-х альтернативная, непараметричная, сложная гипотеза. Она наз-ся гипотезой о законе распределения.Проверяется гипотеза Н0 о том, что генеральная совокупность имеет распределение с плотностью вероятности  против альтернативы с некоторой другой плотностью вероятности. Критерий для проверки таких гипотез наз-ся критерием согласия (это критерий значимости **χ2**). Для проверки рассматриваемой гипотезы используется статистика: (1)Для расчета эмпирического значения статистики v используются данные, такие же, как и при построении гистограммы. Это значит, что весь диапазон выборочных значений разбивается на l интервалов и подсчитывается число выборочных значений, попавших в интервал.mi – число выборочных значений, попавших в i-й интервал, , n – объем выборки, - гипотетическая плотность вероятности попадания СВ ξ в i-й интервал. Она определяется по гипотетической плотности вероятности с помощью формулы:, Δi – i-ый интервал. Поскольку гипотетическая плотность вероятности известна с точностью до параметра, то вместо неизвестных параметров используем их максимально правдоподобные оценки.К. Пирсон доказал, что при увеличении объема выборки n распределение статистики (1) приближается к распределению **χ2**, т.е. HL-m-1. При большом n можно считать, что . Если гипотетическая плотность не имеет параметров, то HL-1. | **30. Проверка гипотезы о равенстве математических ожиданий двух нормальных генеральных совокупностей**Постановка задачи: Имеются 2 выборки разных объёмов из 2х распределений: х1...хm из распределения , y1...yn из норм. генер. совокупн.. Требуется по этим выборкам проверить гипотезу .Решение: будем проверять с помощью значимости.Обозначим -выборочные средние и выборочные дисперсии из этих совокупностей.,, тогда .Должен быть известен закон распределения статистики при Н­0 - истинно.Если Н­0 верна, т.е. , то ,тогда Далее мы знаем, что , , тогда , - с пом. этой статистики провер. гипотеза.Эта статистика не содержит неизвестных параметров и изветен ее закон распределения, след-но она может приниматься для проверки гипотезы.При применяется двухсторонний критерий значимости: .**37. Теория стат. решений с наблюдениями. Случай дискретных состояний и непрерывных наблюдений. Вероятностный смысл среднего риска в случае (0,1) -матрицы потерь.**Рассмотрим следующее выражение:Легко понять, что это выражение представляет собой вероятность , принятие решения ,когда система находится в состоянии .С учетом этого обозначения средний риск примет зн-е:. В случае (0,1)- матрицы потерь выр-е: (\*) Это выражение представляет собой вероятность ошибки (\*)=p(s≠d/). В этом случае средний риск окажется равным r=-средняя (безусловная) вероятность ошибки. Т.о. при (0,1)- матрице потерь оптим. решение обеспечивает min безусловную вероятность ошибки принятия решения. |
| **34.Теория статистических решений без наблюдения****в случае непрерывных состояний и решений**Пусть имеется некоторая система, которая может находится в состоянии , где S – мн-во состояний, кот. предназн. котинуумам. Состояние системы случайные, s-х извест плотностью вероятности. Случ. велич. Имеет плотность распр. f(s). Требуется не производя наблюдений над системой принять решение (указать состояние s). Решение статистики обозн. , где D – пространство решений, аналогичное пространству состояния S.Решение можт быть правильным и не правильным; в случае неправильного решения статистик должен нести потери. Для характеристики потерь вводится функция потерь. Чаще всего функция потерь зависит от рас-ия м/у s и d – q(s,d) Функция потерь должна удовлетворять следующим свойствам:1). W(0)=0 ; 2). W(q) >0; 3). W(q2)≥ W(q2), если q2≥q1. Решение будем рассматривать также случайное, т.е. имеется плотность вероятности f(d). Плотность вероятности f(d) называется решающей функцией, она называется рандомизированной. Поскольку s и d случайные, то функция потерь явл. случайной величиной, следовательно, качество решения характеризуется математическим ожиданием функции потерь, которое называется средним риском r=E(w(s,d)) Задача теории статистических решений формулируется как оптимизационная: требуется найти решающую функцию f(d), минимизирующую средний риск. В виде одного выражения эта задача записывается так: \* Любая задача, в которой предполагаются известными все априорные распределения наз-ся байесовской, а подход такой – байесовский. Применяя байесовский подход запишем выражение среднего риска от мат. ожидания функции потерь. Получим средний риск:, где -дельта функция; -решение, доставляющее минимум условному риску r(d), т.е.  **40. Теория статистических решений с наблюдениями. Случай дискретных состояний и непрерывных наблюдений. Решение задачи в случае (0,1)-матрицы потерь.**Вероятностный смысл среднего риска при (0,1)-матрицы потерь: оптимальное решение обеспечивает минимальную безусловную вероятность ошибки принятия решения. Если выбрать (0,1)-матрицу потерь, то решающее правило можно упростить в сторону уменьшения объема вычислений.Вернемся к случаю 2-х состояний: в этом случае решение можно несколько упростить.  и выберем получим, что Из теории вер-ти изв-на формула полной вер-ти. Получим , что . Сравнивая получим: Теперь след. правило можно записать: . | **35.Теория статистических решений с наблюдениями.****Случай непрерывных состояний, решений и наблюдений.**Система может находиться в состоянии sS и известна плотность вероятности f(s). Над системой выполняются измерения, в рез-те чего получается вектор измерений , где X –пространство измерений. Требуется принять решение о состоянии системы, где D пространство решений. Для того, что бы вектор был получен при принятии решения, он д.б. связан с состоянием системы. Будем считать , что эта связь описывается условной пл-тью вероятности:  и эта пл-ть вер-ти нам известна.В данной задаче решение опред. По наблюдениям. т. е. d=d(x) –рандомизированной решение. Теория статист. решений в общем случае рассматривает рандомизированные решения. В данной задаче рандомизированные решения представляют собой пл-ть вероятности f(d), заданную в пространстве решений d. f(d)- называется решающей функцией. Задача состоит в опред. f(d) -? f(d)=f(d()) минимизируещее среднии риск E = E(w(s,d)). Для характеристики качества решения рассматривается ф-ия потерь и вводится средний риск: r=E(w(s,d)). Задача состоит в том, чтобы минимизировать средний риск: Запишем выражение среднего риска: Минимизация условного риска r(d) эквивалентна минимизации условного риска , получим:.**39. Теория статистических решений с наблюдениями. Случай дискретных состояний и непрерывных наблюдений. Решение задачи в случае двух состояний.**Часто 1 из вариантов выбир.путем парного сравнения. Поэтому рассмотрим ф-цию , где одна из ф-ций определяется через При исп-нии ф-ции решение выносится по правилу: Ф-ция в случае 2-х решений наз-ся дескриминанта или разделяющая ф-ция. Она призвана выбрать одно из 2-х решений.Ур-ние ()=0 определяется в пространстве наблюдений Х( Х) некоторую гиперповерхность, кот.наз-ся разделяющей. Выбор того или иного решения зависит от того, по какую сторону разд-ей гиперповерхности будет находиться наша точка измерения. В случае 2-х измерений =() пр-во наблюд Х предст-т собой плоскость, и гиперпов-ть предст-т собой кривую:Вычисляем распределяющую ф-цию при 2-х составляющих: , L=2, i=1,2, j=1,2. . | **36. Теория стат. решений с наблюдениями. Случай дискретных состояний, дискретных решений и непрерывных наблюдений.**Постановка задачи: Некоторая система может находится в одном из состояний с различными вероятностями Pi=р(Si), . Над системой выполняются наблюдения, в результате чего получаем вектор наблюдений . Этот вектор считается связанным с состояниями системы и эта связь нам известна, задается  (условная плотность вероятности) .Требуется принять о системе одно из решений . Такая задача возникает, например, при стат. распознавании образов. Имеем L образов, каждый характеризуется вектором признаков - требуется по наблюдению этого вектора признаков сказать к какому образу он принадлежит.Будем рассматривать рандомизированные решения, когда на пространстве решений D задано распределение вероятностей Р(dj), j=1,...,n, т.е. каждому решению предписывается вероятность.Имея рандомизированные решения, мы можем в качестве решения dj\* взять решение, имеющее максимальную вероятность.Решение называется нерандомизированным, если распределение вероятностей сконцентрированном в одной точке, если все Р(dj) = 0, кроме одной равной 1.Качество решения будем характеризовать функцией потерь w(s,d) поскольку в данном случае состояния решений дискретны, то функция потерь представляется в виде матрицы потерь .Элементы wij представляют собой потери, которые мы несем в случае, когда принимаем решение dj , а система находится в состоянии Sj. Таким образом, строки матрицы соответствуют состояниям систем, а столбцы – решениям.Элемент характеризует потери в случае правильного решения. На практике чаще всего используется, т. н., “1-0” матрица потерь. Элементы такой матрицы определяются формулой: , где - символ Кронекера.Потери в случае правильного решения равны 0, а в случае неправильного 1. Задача заключается в том, чтобы найти решающую функцию Р(dj), минимизирующую средний риск r=E(w(s,d))– > minВыражение среднего риска: .Где подлеж-я определ-ю вер-ть 0≤Р(dj) ≤1, Σ Р(dj)=1.**42.Статистическое распознавание гаусовских многомерных образов. Постановка задачи, логарифмическая дискриминантная функция.**Данная задача является иллюстрацией теории принятия решений по непрерывным наблюдениям в случае дискретных состояний. Состояния будем называть образом или классом. Предположим, что имеется L-образов. Образы представляються на распознавание с вероятностью . Каждый образ, будем считать, характеризуется n-мерным вектором . - отдельный признак. Значение признака является случайным. Будем считать известными плотности вероятности векторов признаков для каждого образа . Распознавание образов состоит в следующем: измеряется вектор признаков . Требуется сказать к какому классу он принадлежит. Поставленная задача явл. задачей о состоянии и наблюдении, а . Если образы распознаються попарно, то , и применяется правило распознавания: . |
| **38. Теория стат. решений с наблюдениями. Случай дискретных состояний и непрерывных наблюдений. Общее решение задачи.** Постановка задачи: Некоторая система может находится в одном из состояний  с различными вероятностями Pi=р(Si),  Над системой выполняются наблюдения, в результате чего получаем вектор наблюдений . Этот вектор считается связанным с состояниями системы и эта связь нам известна, задается  (условная плотность вероятности) . Требуется принять о системе одно из решений .Такая задача возникает, например, при стат. распознавании образов. Имеем L образов, каждый характеризуется вектором признаков - требуется по наблюдению этого вектора признаков сказать к какому образу он принадлежит.Будем рассматривать рандомизированные решения, когда на пространстве решений D задано распределение вероятностей Р(dj), j=1,...,n, т.е. каждому решению предписывается вероятность. Имея рандомизированные решения, мы можем в качестве решения dj\* взять решение, имеющее максимальную вероятность.Решение называется нерандомизированным, если распределение вероятностей сконцентрированном в одной точке, если все Р(dj) = 0, кроме одной равной 1.Качество решения будем характеризовать функцией потерь w(s,d) поскольку в данном случае состояния решений дискретны, то функция потерь представляется в виде матрицы потерь  .Элементы wij представляют собой потери, которые мы несем в случае, когда принимаем решение dj , а система находится в состоянии Sj. Таким образом, строки матрицы соответствуют состояниям систем, а столбцы – решениям.Элемент wij характеризует потери в случае правильного решения. На практике чаще всего используется, т. н., “1-0” матрица потерь. Элементы такой матрицы определяются формулой:, где  - символ Кронекера.Потери в случае правильного решения равны 0, а в случае неправильного 1.Задача заключается в том, чтобы найти решающую функцию Р(dj), минимизирующую средний риск r=E(w(s,d))– > min Р(dj)Выражение среднего риска: .Где подлеж-я определ-ю вер-ть 0≤Р(dj) ≤1, ΣР(dj)=1.  | **43. Однофакторный дисперсионный анализ. Описание и постановка задачи. Оценка параметров.**Постановка задачи: Для пояснения задачи диспер-ого анализа вернемся к задаче проверки гипотезы о равенстве математ-их ожиданий двух норм. генер-х совокупностей. 2 выборки: ,из 2-х нормальных генер. совокупн.  и . ( и неизвестны).Требуется по этим выборкам проверить гипотезу . Для решения этой задачи используется t-статистика, основанная на сравнении .Задача однофакторного дисп-го анализа явл-ся обобщением этой задачи в направлении нескольких генеральных совокупностей и используется для проверки гипотезы не t-статистики, а f-статистики.Предполагается, что имеется k выборок разных объемов: выборка объемом n1;  выборка объемом n2; … выборка объемом nkиз k нормальных генеральных совокупностей; Задача состоит в проверке гипотезы  против альтернативы, что хотя бы одно из этих равенств не выполняется.Сформулируем эту задачу более коротко: даны наблюдения вида , где ξi,j – ошибки измерений, представляющие собой независимые по i и по j случайные величины с распределением . Требуется по этим данным проверить .Данная задача имеет следующее применение: исследуем влияние некоторого фактора на некоторую характеристику производственного процесса. И фактор может принимать k уровней. Например, пусть исслед-ся влияние стажа работы на производит-ть: В наст. время существует 4 уровня стажа: 1: до 5 лет, 2: от 5 до 10 лет, 3: от 10 до 15 лет, 4: свыше 15 лет.Провер. гипотеза a1=a2=a3=a4 по результатам измер. произв-ти труда большой группы работников с разл. уровнем стажа. Если будет, что Ho принята, то это будет означать, что стаж не влияет на произв-ть. Если отклонение – то влияет. | **41. Проверка простой двухальтернативной гипотезы о математическом ожидании нормальной генеральной совокупности.**Пусть имеется выборка  из нормальной генеральной совокупности , и требуется проверить гипотезу вида , где – значения дискретной случайной величины, имеющие вероятности . Мы видим, что это задача оптимального статистического решения с двумя состояниями , и непрерывными наблюдениями. Для решения задачи выберем (0,1) - матрицу потерь и воспользуемся логарифмическим отношением правдоподобияи решающим правилом:  и мы получим, где .  Решающее правило для нашей задачи имеет видгде  – порог. Предположим, что . Этого всегда можно достичь выбором гипотезы . Тогда наше решающее правило можно представить в виде, где  – новый порог. Мы видим, что решение о математическом ожидании принимается путем сравнения его оценки с некоторым числом . Это число есть сумма середины отрезка ] и порога . Геометрический смысл решающего правила поясняется рисунком, на котором изображен отрезок действительной прямой и его середина – точка . Решение выносится в пользу правой границы отрезка, если выборочное среднее превышает середину отрезка плюс порог . Рис6_1 |
|  |  |  |