定义2.1.1 条件概率 假设我们已知事件B发生并且我们想要计算另一个事件A在事件B已经发生条件下的概率。新的事件A概率则被称为事件A在事件B已经发生条件下的条件概率(conditional probabilty of the event A given that the event B has occurred)，标记为.如果>0，则条件概率为

 (2.1.1)

当=0时，条件概率未定义。

定理2.1.1 条件概率的乘法法则 设A和B是事件。如果>0, 则

.

如果,则

.

定理2.1.2 条件概率的乘法法则 假设事件,...,满足则

 (2.1.2)

条件概率和分割

定义2.1.2 分割 设S为一次实验的样本空间，考虑S中k个不相交事件并且那么这些事件构成了S的一个分割。

分割能促进某些事件概率的计算.

定理2.1.4 全概率公式 假设事件构成空间S的一个分割并且对于任意成立. 那么对于S中的每一个事件A，存在

 (2.1.4)

全概率公式的条件概率：

 (2.1.5)

**增强实验**. 这个概念有点难理解，需要结合中文教材。

定义2.2.1 独立事件 两个事件A和B是独立的当且



假设,那么很容易从独立事件和条件概率的定义得出：

事件A和事件B是独立的当且仅当并且

定理 2.2.1 如果两个事件A和B是独立的，则事件A与,与B,与同样也是独立事件。

定义2.2.2 (互相)独立事件. k个事件是独立的（或相互独立的），对于这些事件的任意子集()，存在

.

贝叶斯定理：设事件是构成空间S的一个分割并且,其中设A是一个事件且则

 (2.3.1)

条件贝叶斯定理：

 (2.3.2)

定义3.1.1 随机变量 设S是一项实现的样本空间。定义在S上的实值函数被称为随机变量。

3.4 二维分布 *2019年6月4日09点51分*

**定义3.4.1 联合/二维分布** 设和是随机变量. 和的联合分布或二维分布是形如所有概率的集合, 对于中所有实数对都是事件。

**定义3.4.2 离散联合分布** 设和是随机变量, 并考虑有序对(). 如果()对数有限或者可数，则我们说和拥有离散联合分布.

**定理3.4.1** 假设两个随机变量和均是离散分布. 则和拥有离散联合分布.

**定义3.4.3** 联合概率函数 p.f. 和的联合概率函数或联合p.f.定义为函数f使得平面上每一个点,

**定理3.4.2** 设是离散联合分布. 如果不是对的可能值，则, 并且

对于C中的每一个有序对, 则

**定义3.4.4** **连续联合分布/联合p.d.f/支持** 两个随机变量和拥有连续联合分布，如果存在一个非负函数定义在整个平面上，使得平面中每一个子集满足，

如果该积分存在. 则函数被称为的联合概率密度函数(或者联合p.d.f.)。 闭集被称为的(分布)支持.

**定理 3.4.3** 一个联合p.d.f.必须满足下列两个条件：

对于

且

**定理3.4.4** 对于每一个在平面上的连续联合分布，下面两个定理成立:

* + 1. 平面上的每一个点，每一组无穷点序列，其概率可为0.
    2. 设是定义在上的一维连续实函数. 则集合和的概率为0.

**定义3.4.5** **联合p.f./p.d.f.** 设和是随机变量，其中是离散的，是连续的. 假设存在一个函数定义在平面上，对于实数子集中的每一对A和B，

如果该积分存在. 则函数被称为和的联合p.f./p.d.f.

定义3.4.5中和互换也同样成立。每一个联合p.f./p.d.f都必须满足两个条件。如果是离散随机变量并且其可能值为并且是连续随机变量，对于所有,则

定义 联合(累加)分布函数c.d.f. 随机变量和的联合分布函数或联合累加分布函数(联合c.d.f)定义为函数F使得和的所有值(),

如果任意随机变量和的联合c.d.f.是函数F, 那么位于平面上的变量对的概率则位于特殊的矩形区域内，对于给定的数字和，

**定理3.4.5** 如果和拥有联合c.d.f.则单个随机变量的c.d.f. 可以导出得到相似的，的.

**3.5 边际分布**

**定义3.5.1** **边际c.d.f/p.f./p.d.f.** 假设和具有联合分布. 通过定理3.4.5导出的c.d.f被称为的边际c.d.f. 同样，与边际c.d.f.相关联的p.f.或p.d.f.被称为的边际p.f.或p.d.f.

**定理3.5.1** 如果和拥有联合分布p.f.*f*，则的边际p.f.为

相似地，的边际p.f.为

**定理3.5.2** 如果和拥有连续联合分布*f*，则的边际p.d.f.为

相似地，的边际p.f.为

**定理3.5.3** 设f是和的联合p.f./p.d.f, 其中是离散的，是连续的. 则的边际p.f.

并且的边际p.f.为

**定义3.5.2** **独立随机变量** 如果随机变量和是独立的，仅当实数子集和满足(是事件)

**定理3.5.4** 设和拥有联合c.d.f., 设的边际c.d.f.为，设的边际c.d.f.为. 则和是独立的当且仅当对于所有的实数和

**定理3.5.5** 设和为随机变量并且拥有联合p.f.*f*. 则设和是独立的当且仅当*f*可以用下列公式表达

其中是单独关于的函数，是单独关于的函数.

**推论3.5.1** 两个随机变量和是独立的当且仅当所有实数和满足下列因式:

**定理3.5.6** 设和拥有连续联合分布. 假设是一个矩形区域(可能无界限)，边界线与坐标轴(可能)平行. 则和是独立的当且仅当公式(3.5.7)在上成立.

3.6 条件分布 2019年6月17日11点50分

**定义3.6.1** **条件分布/p.f.** 设和拥有离散联合分布*f*. 设为的边界p.f. 对于每一个值使得，定义为

则被称为给定情况下的条件p.f. 相似的，如果给定，则的条件p.f.为

**定义3.6.2** **条件p.d.f.** 设和拥有连续联合分布，其p.d.f.为*f*，相应的边际p.d.f.为和. 对于每一个值使得. 则在给定条件下的条件p.d.f.定义如下：

当值使得时，我们可以自由定义，使得作为的p.d.f.函数。

**定理3.6.1** 对于每一个, 定义3.6.2得到的是的一个p.d.f.函数。即

根据公式(3.5.3).

**定义3.6.3** **混合分布的条件p.f.或p.d.f.** 设是离散的，是连续的并且p.f./p.d.f.为f. 则给定条件下的条件p.f.由公式(3.6.2)定义, 在给定条件下的条件p.f.由公式(3.6.3)定义.

**定理3.6.2** **分布的乘法法则** 设和是随机变量，它们对应的边际p.f./p.d.f.分别为和. 假设在给定条件下的条件p.f.或p.d.f.为，在给定条件下的条件p.f.或p.d.f.为. 则对于每一个值使得和每一个,

其中f为和的联合p.f./p.d.f. 类似的，对于每一个值使得和每一个,

**定理3.6.3** **随机变量的总概率定律** 如果是随机变量的边际p.f.或p.d.f., 是给定条件下的条件p.f.或p.d.f. 则的边际p.f.或p.d.f.为

如果上式中是离散的。如果是连选的，则的边际p.f.或p.d.f.为

**定理3.6.4** **随机变量的贝叶斯定理** 如果是随机变量的边际p.f.或p.d.f.并且是给定条件下的条件p.f.或p.d.f. 则在给定条件下的条件p.f.或p.d.f.为

其中是从公式(3.6.11)或(3.6.12)获取。类似的，给定条件下的条件p.f.或p.d.f.为

**定理3.6.5** **独立随机变量** 假设和是两个随机变量并且它们的联合p.f./p.d.f.为*f.*则和是独立的当且仅当对于每一个值使得和每一个, 满足下列公式

3.7 多变量分布 2019年6月19日16点48分

**定义3.7.1 联合分布函数/c.d.f.** 多变量的联合c.d.f.用表示，其值分布在维空间各个点()上，存在如下关系：

每一个多变量c.d.f.都满足之前单变量和二元变量c.d.f.的定义的属性。

**定义3.7.2** **联合离散分布/p.f.** 如果随机向量()拥有有限值或者是中()不同可能值的无穷序列，那么个随机变量拥有离散联合分布。则的联合p.f.定义为函数使得每一个点,

用向量标记

**定理3.7.1** 如果是联合离散分布并且联合p.f.为, 则对于每一个子集

**定义3.7.3** **连续分布/p.d.f.** 如果在上存在一个非负函数使得对于每一个子集满足

则称这个随机变量拥有连续联合分布。函数被称为的联合p.d.f.

用向量标记

**定理3.7.2** 如果的联合分布是连续的，则联合p.d.f.函数可以从联合c.d.f.函数导出

**推导边际p.d.f.** 如果个随机变量的联合分布是已知的，那么单个随机变量的边际分布能够从它们的联合分布导出。例如，如果的联合p.d.f.为, 则的边际分布在每个值处为

更一般的说，个随机变量中任意个联合边际p.d.f.可以通过在其余个变量对联合p.d.f.作积分获得。例如，设四个随机变量和的联合p.d.f.为，则和的边际二元p.d.f.在每个点()定于为

**推到边际c.d.f.** 如果个随机变量的联合c.d.f.为. 则的边际c.d.f.可以通过下列关系获得：

**定义3.7.5** 如果实数的任意个集合满足下列等式

则个随机变量是独立的。

**定理3.7.3** 设随机变量的联合c.d.f.为F, 的边际单变量c.d.f.为，其中. 则随机变量是独立的当且仅当，对于所有点，满足

**定理3.7.4** 如果随机变量是离散的，连续的，或者是混合的。设它们的联合p.d.f.,联合p.f.或联合p.f./p.d.f.为f, 并且是的边际单变量p.d.f.或p.f., 则随机变量是独立的当且仅当，对于所有点，满足

**定义3.7.6** **随机样本/i.i.d./样本大小**. 考虑实线上一个给定的概率分布，用p.f.或p.d.f. 表示。如果这些随机变量是独立的并且它们的边际p.f.或p.d.f.为，那么这个随机变量构成该分布的一个随机样本。这样的随机变量也被称为独立同分布(independent and identically distributed)，简称i.i.d. 我们把随机变量的数目称为样本大小。

定义3.7.6称构成分布f的一个样本空间，如果它们的联合p.f.或p.d.f. g对于所有点满足如下等式:

显然，i.i.d.样本不能有混合分布。

**定义3.7.7** 条件p.f.,p.d.f,或p.f./p.d.f. 假设随机变量被分割成两个子向量和, 其中是一个维随机向量，包含中个随机变量，是一个维随机向量包含另外个随机变量。假设的n维联合p.f.,p.d.f.,或p.f./p.d.f.是并且的边际维p.f.,p.d.f.,或p.f./p.d.f为. 对于给定的使得，则在给定条件下的条件维p.f.,p.d.f.,或p.f./p.d.f. 定义如下：

公式(3.7.8)可以被写为

**定理3.7.5** **多变量总概率定律和贝叶斯理论** 假设条件和符号如定义3.7.7。如果拥有连续联合分布，则的边际p.d.f.为

并且在给定条件下的条件p.d.f为

**定义3.7.8 条件独立随机变量** 设是一个随机向量并且它的联合p.f.,p.d.f.,或p.f./p.d.f.为. 多个随机变量在给定的情况下是条件独立的，当且对于所有值使得,满足

其中代表在给定条件下的条件多变量p.f.,p.d.f.,或p.f./p.d.f. 代表在给定条件下条件变量p.f或,p.d.f.

**定义3.7.9 直方图.** 设是一组位于之间的数字，. 选择某些整数，将区间分割成长度为的k个等长子区间。对于每一个子区间，统计有多少数字在该子区间。设为子区间的个数, . 选择一个数字(一般或或). 绘制一个二维图像，其横轴范围从到. 对每一个子区间绘制一个长为，高为，横过第个区间的矩形条。这样的图像被称为直方图。

3.8 随机变量函数

**定理3.8.1 离散随机变量函数** 设是离散分布，其p.f.为，设是定义在所有值集合上某个函数。对于的每一个值，的p.f. g为

一般的，假设的p.d.f.为，另一个随机变量定义为。对于每一个实数，的c.d.f. 可以从下列公式导出：

如果随机变量同样拥有连续分布，它的p.d.f. 可以由下列公式获得：

**定理3.8.2 线性函数**. 假设是一个随机变量，其p.d.f.为，设 则的p.d.f.为

**定理3.8.3 概率积分变换** 设的连续c.d.f为，设. (从的变换被称为概率积分变换。)则的分布是区间[0,1]上的均匀分布。

**推论3.8.1** 设在[0,1]区间上是均匀分布，设是一个连续的c.d.f并且分位函数为. 则拥有c.d.f. F.

**定理3.8.4** 设是一个随机变量，其p.d.f.为， (这里和/或既可以是有限的也可以是无限的。)设并假设是可微的并且在是一对一的。设是区间在函数映射后的图像。设是在区间上的反函数。则的p.d.f. 为

3.9 二元或多元随机变量函数 2019年7月1日09点21分

**定理3.9.1 离散随机变量函数** 假设个随机变量拥有联合分布，其联合p.f.为，由这个随机变量构成的个函数定义如下：

为这个函数随机变量给定一些初始值，用标记全部点的集合使得

的联合p.f. g在特定点是通过关系

**定理3.9.2 二项和伯努利分布** 假设是i.i.d随机变量并拥有参数为p的伯努利分布。设。则拥有参数为和p的二项分布。

**定理3.9.3 蛮力分布** 假设的联合p.d.f.为，设. 对于每一个实数，定义。则Y的c.d.f G(y)为

**定理3.9.4 两个随机变量的线性函数** 设拥有联合p.d.f. ，设并且. 则拥有连续分布，其p.d.f.为

**定义3.9.1 卷积** 设是独立随机变量，设. 的分布被称为分布的卷积。的p.d.f.有时被称为的p.d.f.的卷积。

**定理3.9.5 多变量变换** 设拥有连续联合分布，其联合p.d.f.为. 假设存在一个子集S使得. 定义n个新的随机变量如下：

其中我们假设这n个函数定义了S到子集T的变换，并且是一对一可微的。设其逆变换如下定义：

则的联合p.d.f.为

其中的行列式为

表示行列式的绝对值。

**定理3.9.6 线性变换** 设拥有连续联合分布，其联合p.d.f.为. 定义通过下列变换

其中是一个非奇异矩阵。则拥有连续联合分布，其p.d.f.为

3.10 马尔科夫链 2019年7月2日10点26分

**定义3.10.1 随机过程** 随机变量组成的序列被称为离散时间参数的随机过程(stochastic process, random process)。第一个随机变量被称为过程的初始状态；当时，随机变量被称为在时间的过程状态。

**定义3.10.2 马尔科夫链** 一个离散时间参数的随机过程被称为马尔科夫链，当且，对于每一个时间n,在给定条件下，所有的条件分布仅依赖而不是早期状态。用符号表示如下所示：

**定理3.10.1**对于一个有限马尔科夫链，前n个状态的联合p.f.为

同样，对于每一个n和m>0，

**定义3.10.3 过渡分布/平稳过渡分布** 考虑一个由k个状态的马尔科夫链。在时间n+1状态的条件分布是根据在时间n上的状态给出的，也就是，其中并且 ，被称为马尔科夫链的过渡分布。如果过渡分布在每一次时间点n的值都是一样的，则该马尔科夫链拥有平稳过渡分布。用公式表示为

**定义3.10.4 过渡矩阵** 考虑一个平稳过渡分布的有限马尔科夫链，其概率分布为对于所有的. 马尔科夫链的过渡矩阵被定义成矩阵，也就是

列表示当前状态，行预测下一状态。

**定义3.10.5 随机矩阵** 如果一个方阵的所有元素非负并且每一行的元素之和为1，则该矩阵被称为随机矩阵。

**定理3.10.2 多步过渡** 设是一个平稳过渡分布有限马尔科夫链的过渡矩阵。对于每一个，矩阵的第个指数中第行第列的概率表示从状态到状态的步转移。(**这里应该是从状态到状态的步转移**)

**定义3.10.6 多布过渡矩阵** 在定理3.10.2条件下，矩阵被称为马尔科夫链的步*过渡矩阵*。

**定义3.10.7 吸收状态** 在马尔科夫链中，如果对于某些状态，则这种状态被称为吸收状态。

**定义3.10.8** 一个包含非负元素且总和为1的向量被称为概率向量。概率向量被称为链的初始分布或初始概率向量，其坐标设定马尔科夫链将要进入每一个时间1状态时的概率。

初始分布和过渡矩阵共同决定了马尔科夫链的联合分布。设为厨师分布，公式(3.10.1)则被重写为

**定理3.10.3 时间n的边际分布** 考虑一个平稳过渡分布的有限马尔科夫链，其初始分布为,过渡矩阵为。则在时间n的边际分布为.(**证明过程每看懂**)

**定义3.10.9 平稳分布** 设是马尔科夫链的一个过渡矩阵。概率向量满足被称为马尔科夫链的平稳分布。

**定理3.10.4** 如果存在使得每一个元素都为正数，则

* 该马尔科夫链存在唯一的平稳分布,
* 是一个矩阵，其所有的行都等于，并且
* 不论马尔科夫链以哪种分布开始，其分布经过步过渡之后随收敛于.

**4.1 随机变量的期望** 2019年7月3日09点43分

**定义4.1.1 有界离散随机变量的均值** 设是一个有界离散随机变量，其p.f.为。 的期望，标记未，定义如下：

的期望也被称为的均值或的期望值.

**定义4.1.2 一般随机变量的均值** 设是一个随机变量，其p.d.f.为。假设下列之至少有一个是有限的：

则的均值或期望存在，定义如下

如果(4.1.2)中的两个和式都是无限的，则不存在。

**定义4.1.3 有界连续随机变量均值** 设是一个有界连续随机变量，其p.d.f.为。的期望标记为，定义为:

**定义4.1.4 一般有界连续随机变量均值** 设是一个有界连续随机变量，其p.d.f.为。假设下列积分至少有一个是有限的：

则的均值或期望存在，定义如下

**定理4.1.1 佚名统计学家定理(Law of unconscious statistician)** 设是一个随机变量，设是一个实变量的实值函数。如果是连续分布并且均值存在，则

如果是离散分布且均值存在，则

**定理4.1.2 佚名统计学家定理(Law of unconscious statistician)** 假设是随机变量联合p.d.f.为并且存在均值。设r是n个实变量的实值函数，假设。则可以从下列关系中直接得出

相似地，如果是离散联合分布，p.f.为并且存在均值，则的均值为

4.2 期望的属性 2019年7月8日09点10分

**定理4.2.1 线性函数** 如果，其中和都是有限常量，则

**推论 4.2.1** 如果且概率为1，则.

**定理4.2.2** 如果存在一个常数使得，则. 如果存在一个常数b使得，则。

**定理4.2.3** 假设，并且或. 则

**定理4.2.4** 如果是n个随机变量使得每一个期望是有限的()，则

**推论4.2.2.** 假设期望是有限的(). 所有常数和b存在下列关系:

**定义4.2.1** 凸函数 一个以向量为参数的函数是凸的，当且对于每一个,和每一个，,满足下列不等式

**定理4.2.5 Jensen’s不等式** 设是一个凸函数，设是一个有限均值的随机变量. 则

**定理4.2.6** 如果是n个独立随机变量使得每一个期望是有限的()，则

**定理4.2.7 整型随机变量** 设是一个随机变量只能取值则

**定理4.2.8 一般非负随机变量** 设是一个非负随机变量，其c.d.f.为. 则

4.3 方差 2019年7月9日09点34分

**定义4.3.1 方差/标准偏差** 设是一个随机变量，其有限均值. 的方差，标记为，定义如下：

如果具有无限均值或者如果的均值不存在，我们说不存在. 如果方差存在，则的标准偏差是的非负平方根。

当仅讨论一个随机变量时，通常用符号表示其标准偏差，方差用表示. 当讨论多个随机变量时，随机变量的名称作为符号的下标包括在内，例如，将是的标准偏差，而将是的方差。

**注意：方差仅取决于分布.** 随机变量的方差和标准差仅取决于的分布，正如的期望仅取决于分布。 事实上，所有可以从p.f. 或者p.d.f.计算的东西仅取决于分布。具有相同分布的两个随机变量将具有相同的方差，即使它们彼此无关。

**定理4.3.1 计算方差的另一种方法** 对于每一个随机变量，.

方差（以及标准偏差）提供了围绕其均值分布的扩散或分散的确定性。较小的方差值表示概率分布紧密集中在附近; 较大的方差值通常表示概率分布在附近具有较宽的范围。然而，将非常小但正的概率放置在距离实线上原点足够远的地方，可以使分布的方差及其均值任意大。

**定理4.3.2** 对于每一个, . 如果是一个有界随机变量，则必定存在并且是有限的。

**定理4.3.3** 当且仅当存在一个常数c使得.

**定理4.3.4** 对于常数和, 设. 则

且

**定理4.3.5** 如果是n个独立随机变量且包含有限均值，则

**推论4.3.1** 如果是n个独立随机变量且包含有限均值，如果和b是任意常数，则

**定义4.3.2 四分位距(Interquartile Range IQR)** 设是一个随机变量且分位函数为. 则四分位距(IQR)定义为.

换句话说，IQR是包含分布中间一半长度的区间。