OLASILIK TEORİSİ

Hatırlanacağı gibi, ilk bölümde İstatistiğin Betimsel ve Tümevarımsal İstatistik diye iki kısma ayrıldığını söylemiştik. Tümevarımsal İstatistik, örnekten hareketle yığına ilişkin bilinmeyen parametreler hakkında tahminlerde bulunmak amacıyla geliştirilmiş yöntemler bütünüdür. Tümevarımsal istatistik yöntemleri olasılık teorisinin gelişmesiyle ortaya çıkmıştır. Bu nedenle, bundan sonraki bölümlerin ağırlıklı konusu olan tümevarımsal istastistik yöntemlerine geçmeden önce olasılık ile ilgili bilgilerimizi tazeleyelim.

Olasılığın tanımı ve türleri anlatılırken kullanacağımız faktöriyel notasyonu ile küme ve olay kavramlarına bir göz atmak gerekir. çünkü, olasılık kavramları küme ve olay kavramları üzerinde yükselmektedir. 1'den n'e kadar sayıların çarpımı faktöriyel notasyonu ile $n!$ şeklinde yazılır ve "n faktöriyel" diye okunur. Yani,

$$n! = 1.2.3...(n-2).(n-1).n$$

olarak yazılır. Örneğin,

$$3! = 3.2.1 = 6$$

olurken

$$5! = 5.4.3.2.1 = 120$$

olur. Faktöriyellerin bölümünde sadeleştirme yapılabilir. Örneğin;

$$\frac{5!}{3!} = \frac{5.4.\cancel{3}.\cancel{2}.\cancel{1}}{\cancel{3}.\cancel{2}.\cancel{1}} = 5.4 = 20$$

Bu arada hemen belirtelim ki tanım olarak,

$$0! = 1! = 1$$

olduğu kabul edilir.

8.1. Permütasyon ve Kombinasyon

Herhangi bir deneyde sonuçların olasılık hesabı yapılırken Elverişli Sonuçların Sayısı (ESS) ve Mümkün tüm Sonuçların Sayısının (MSS) bilinmesi gerekir. bu sayıların elde edilmesinde permütasyon ve kombinasyondan yararlanılır. Permütasyon, n elemanın r li olarak kaç farklı şekilde sıralanabileceğine cevap verir. Permütasyon sayıları bulunurken bir elemana sıraya girmesi için birden fazla şans verilebilir veya verilmez. Eğer sıraya giren bir elemana sonraki bir sırada tekrar şans verilirse tekrarlı permütasyon durumu ortaya çıkar. Örneğin A, B, C, D harflerinden ikişerli kelimeler oluştururken harflerin tekrar etmesine izin veriyorsak $4.4 = 16$ diziliş söz konusu olur. bu 16 diziliş;

AA	AB	AC	AD
BA	BB	BC	BD
CA	CB	CC	CD
DA	DB	DC	DD

şeklinde gösterilebilir. Genel olarak *n* sayıda elemanın *r* elemanlı tekrarlı permütasyonu;

$$P\begin{pmatrix} n \\ r \end{pmatrix} = n^r$$

ile verilir.

Örnek : Bir paranın arka arkaya üç kez atılışında ortaya çıkan sonuçların sayısının kaç olduğuna bakalım. Burada iki elemanın (yazı, tura) üçlü permütasyonu söz konusudur. Ve bu tekrarlı bir permütasyondur. Çünkü birinci atışta örneğin yazı gelse, ikinci hatta üçüncü atışta yine yazı gelebilir. O halde burada $n = 2$ ve $r = 3$ olmak üzere $n^t = 2^3 = 8$ mümkün durum vardır. Bu durumlar;

YYY	YYT	YTY	YTT
TTT	TTY	TYT	TYY

şeklinde liste olarak görülebilir. Permütasyonda eğer bir elemana birden fazla yer verilmiyorsa bu kez tekrarsız permütasyon söz konusu olur. *n* elemanın *r* li tekrarsız permütasyonu;

$$P\begin{pmatrix} n \\ r \end{pmatrix} = \frac{n!}{(n-r)!}$$

ile hesaplanır.

Örnek : A, B, C, D harfleri tekrarsız ve ikişerli olmak üzere kaç farklı şekilde dizilebilir?

Çözüm: Burada 4 elemanın 2'li permütasyonu söz konusudur. Yani;

$$P\begin{pmatrix} 4 \\ 2 \end{pmatrix} = \frac{4!}{(4-2)!} = \frac{4.3.\cancel{2}.\cancel{1}}{\cancel{2}.\cancel{1}} = 12$$

Bu 12 diziliş aşağıda gösterilmiştir.

-	AB	AC	AD
BA	-	BC	BD
CA	CB	-	CD
DA	DB	DC	-

Görüldüğü gibi tekrarlı durumların yer aldığı köşegen elemanlarına yer verilmemiştir. Eğer *n* elemanın *n*'li permütasyonunu arıyorsak;

$$P\binom{n}{r} = \frac{n!}{(n-r)!} = \frac{n!}{0!} = n!$$

olarak bulunur.

Örnek : A, B, C, D harfleri ile 4 harfli anlamlı/anlamsız kaç kelime yazılabileceği sorulsa idi;

$$P\binom{4}{4} = \frac{4!}{(4-4)!} = \frac{4!}{0!}$$

$$= 4! = 4.3.2.1 = 24$$

olarak bulunurdu. Kombinasyon ise *n* tane nesne arasından *r* tanesinin kaç farklı şekilde seçilebileceğine cevap verir. Kombinasyonda çekiliş sırası önemli değildir. Buna göre *n*'in *r*'li kombinasyonu;

$$C\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

ile hesaplanmaktadır.

Örnek 4: 5 kişilik bir gruptan 3 kişilik bir heyeti kaç farklı şekilde seçmek mümkündür?

Çözüm: Seçilen şahısların seçiliş sırası önemli olmadığından bu soruya kombinasyon ile cevap aranır. Buna göre;

$$C\binom{5}{3} = \frac{5!}{3!(5-3)!} = \frac{5.4}{2.1} = 10$$

10 farklı şekilde söz konusu heyet teşkil edilebilir. Kombinasyonda bir elemana seçime girmesi için tekrar şansı verilirse tekrarlı kombinasyon ortaya çıkar. *n* elemanın *r*'li tekrarlı kombinasyon sayısı;

$$C\binom{n+r-1}{r}=\frac{(n+r-1)!}{r!(n-1)!}$$

ile hesaplanır.

Örnek : A, B, C, D harfleri verilmişken bu elemanların ikişerli tekrarlı kombinasyonu;

$$C\binom{4+2-1}{2}=\frac{(6-1)!}{2!(4-1)!}=\frac{5!}{2!3!}=10$$

olarak hesaplanır. Bu 10 kombinasyonu;

AA	AB	AC	AD
-	BB	BC	BD
-	-	CC	CD
-	-	-	DD

biçiminde liste halinde yazabiliriz.

Kümeler ve Olaylar

Küme, nesnel ya da kavramsal belirli öğelerden meydana gelen bir topluluktur. Küme elamanlarının aynı türden olması şart değildir. Ancak, elemanlarının belirli olması şarttır. Örneğin;

- Sınıfımızda adı A harfi ile başlayan öğrenciler topluluğu bir kümedir. Belirli elemanlardan oluşmaktadır. Somut bir kümedir. Kümenin elemanları bir araya getirilebilir.
- Sınıfımızda okutulan dersler topluluğu yine bir kümedir. Ancak kavramsal (soyut) bir kümedir. Bu kümenin elemanları nesnel olarak bir araya getirilemez fakat simgesel olarak temsil edilebilir.
- Sınıfımızdaki zeki öğrenciler ya da dünyadaki dürüst insanlar topluluğu, zekânın ve dürüstlüğün nesnel ölçüleri olmadığı sürece bunlar birer küme olarak kabul edilemez. Çünkü hangi elemanlardan meydana geldiği belli değildir. Küme elemanları liste ya da tanım halinde gösterilebilir. Örneğin 0 ile 100 arasındaki tek sayılardan oluşan küme,

Liste halinde $A=\{1, 3, 5, 7, 9,\ldots,97, 99\}$

Tanım olarak $A=\{x|\ 0<x<100,\ \ x$ tek sayı, $x\in N\}$ ile gösterilebilir.

Bunun dışında, kümeleri Venn diyagramı denilen şekillerle de gösterebiliriz. Hiç elemanı olmayan kümeye boş küme denir ve $\varnothing$ ya da $\{\ \}$ ile gösterilir. Herhangi bir A kümesinin bir a elemanı, $a\in A$ (a eleman A) diye yazılır. A nın elemanı olmayan bir b nesnesi söz konusu ise $b\notin A$ (b A'nın elemanı değil) şeklinde gösterilir. Belli bir tanıma göre bir araya gelen bütün kümelerin oluşturduğu topluluğa Evrensel Küme (E kümesi) diyoruz. Bir kümenin elemanları ile daha az sayıda eleman içeren farklı farklı küçük kümeler oluşturulabilir ki bu kümelere alt kümeler adı verilir. Şimdi kümeler üzerindeki bazı işlemleri simge olarak verelim. A ve B herhangi iki küme olsun.

$A\subset B$	*A*, *B*'nin öz alt kümesi
$A\not\subset B$	*A*, *B*'nin alt kümesi değil
$A\subseteq B$	*A*, *B*'nin alt kümesi (kümeler eşit de olabilir)
A'	*A*'nın değili kümesi
$A\cup B$	*A* ve *B* kümelerinin bileşimi kümesi
$A\cap B$	*A* ve *B*'nin kesişimi kümesi
$A-B$	*A*'da olan fakat *B*'de olmayan elemanların oluşturduğu küme

A ve B gibi iki kümenin hiçbir ortak elemanı yoksa, bu iki kümeye ayrık kümeler denir. bir deneysel süreçte ortaya çıkabilecek her sonuç bir küme, bütün sonuçların bileşimiyle de örnek uzayı (evrensel küme) oluşur. Örnek uzayının her alt kümesi bir olay olarak tanımlanabilir. İstatistikte yığın ve örnek tipik birer küme kavramıdır. Yığın kümesi, örnek adı verilen alt kümelerden oluşur. Başka bir deyişle örneklerin oluşturduğu örnek uzayına yani evrensel kümeye yığın (anakitle, kitle) adı verilir.

Örnek : Hilsesiz bir zarın atılması deneyini düşünün. Bu atılışta 1, 2, 3, 4, 5 veya 6 gelmesi sonuçlarının her biri bir kümedir. Bu kümeler topyekün bir evrensel küme oluşturur. Bu evrensel kümenin çeşitli alt kümeleri düşünülebilir. Buna göre söz gelimi

$A=\{1\}$	A olayı: zarın 1 gelmesi
$B=\{1, 3, 5\}$	B olayı: zarın tek sayı gelmesi

Gibi olaylar tanımlanabilir. Sadece bir sonucu içeren olaylara basit olay, birden fazla sonucu içeren olaylara ise bileşik olay denir. yukarıdaki örnekte A basit olay olurken, B bileşik bir olaydır. İki olay bir deneyde aynı anda meydana gelemiyorsa bunlara ayrık olaylar denir. zar atma deneyinde 1, 2, 3, 4, 5 veya 6 gelmesi olaylarının hepsi karşılıklı ayrık olaylardır. Aynı anda bu olaylardan ikisi ya da daha fazlası meydana gelemez. İki olaydan birinin sonucu diğerini etkilemiyorsa bu olaylar birbirinden bağımsız olaylardır. Eğer iki olaydan birinin olması diğer olayın meydana gelmesini etkiliyorsa bunlara bağımlı olaylar denir. İçerisinde 3 kırmızı 4 beyaz top bulunan torbadan iadeli olarak iki top çektiğimizde birinci topun kırmızı olması olayı ile ikinci topun kırmızı olması olayları birbirini etkilemez. Bu iki olay bağımsızdır. Torbalar iadesiz çekildiğinde bu iki olay bağımlı hale gelir. Öte yandan, bir boş kümeye imkânsız olay, evrensel kümeye ise kesin olay denilebilir. Zar atma deneyinde 7 gelmesi kümesinin hiçbir elemanı yoktur. Bu bir boş kümedir ve imkânsız olaydır. Zarın bir atılışında evrensel küme 1, 2, 3, 4, 5 veya 6 sonuçlarından oluşur. Zarın atılışında bu kümenin elemanlarından birinin gelmesi kesindir. Yani evrensel küme kesin olaydır.

Olasılık ve Türleri

Bir deneyde herhangi bir olayın olasılığı, söz konusu olayın olması için elverişli sonuçların sayısı, mümkün tüm sonuçların sayısına bölünerek elde edilir. Bir A olayının olasılığı $P(A)$ ile gösterilir. Buna göre;

$$P(A) = \frac{ESS}{MSS}$$

olur. ESS sıfırdan küçük, MSS'den de büyük olmayacağına göre, $P(A)$ da, sıfırdan küçük 1'den büyük olamaz. Yani

$$0 < ESS < MSS$$

dir. Buna göre, herhangi bir A olayının olasılığı 0 ile 1 arasında $\left(0 \leq P(A) \leq 1\right)$ yüzdelik bir rakam olarak karşımıza çıkar.

Örnek : Zar atma deneyinde, A olayı: çift sayı gelmesi olsun. Bunun için 2, 4 ve 6 sonuçları olmak üzere elverişli 3 sonuç vardır. mümkün tüm sonuçların sayısı ise 6'dır. O halde;

$$P(A) = \frac{ESS}{MSS} \Rightarrow P(A) = \frac{3}{6} = \frac{1}{2}$$

olarak bulunur.

Örnek : Hilesiz madeni bir paranın iki defa atılması deneyini göz önüne alalım. Burada örnek uzayı, yani mümkün tüm haller;

$$E = \{YY,\ YT,\ \ TY,\ TT\}$$

olup bu örnek uzayında bir *C* olayı "iki yazı gelmesi biçiminde verilsin. Buna göre bu *C* olayının olması için elverişli bir sonuç vardır (*YY* sonucu). Mümkün tüm sonuçların sayısı ise 4'tür. Buradan *C*'nin olasılığı;

$$P(C) = \frac{1}{4}$$

olarak bulunur.

A ve *B* olayları verilmişken, (*A* veya *B*) olayı, $A \cup B$ ile (*A* ve *B*) olayı, $A \cap B$ ile gösterilir. (*A* veya *B*) olayının olasılığı; $P(A \cup B)$ ile (*A* ve *B*) olayının olasılığı, $P(A \cap B)$ ile gösterilir. Bağımlı olaylar durumunda olaylardan birinin olması şartı altında diğerinin olması olasılığı, şartlı olasılık adını alır. *A* ve *B* bağımlı olaylar olsun. *B* olayının gerçekleşmiş olması halinde *A*'nın olasılığı (şartlı olasılık), $P(A|B)$ ile gösterilir ve

$$P(A|B)=\frac{P(A\cap B)}{P(B)}$$

şeklinde hesaplanır. Burada $P(A\cap B)$, *A* ve *B*'nin aynı anda meydana gelme olasılığıdır. Eğer *A* ve *B* olayları bağımsız ise;

$$P(A|B)=P(A)$$

olur. Yani *B*'nin gerçekleşmiş olması şartı *A*'nın olasılığını etkilemez.

Örnek : Bir torbada 3'ü beyaz 2'si siyah 5 küçük top, 4'ü beyaz, 2'si siyah 6 büyük top vardır. Bu torbadan rastgele bir top çekiliyor. Çekilen topun küçük olduğu bilindiğinde bunun beyaz olması olasılığı nedir?

Çözüm:

A olayı : Çekilen topun beyaz olması

B olayı : Çekilen topun küçük bir top olması

$A\cap B$ olayı : Çekilen top hem beyaz hem de küçük olsun.

$$P(A\cap B)=\frac{3}{11} \qquad P(B)=\frac{5}{11}$$

Buradan şartlı olasılık;

$$P(A|B)=\frac{P(A\cap B)}{P(B)}=\frac{\frac{3}{11}}{\frac{5}{11}}$$

$$P(A|B)=\frac{3}{5}$$

olarak bulunur. Burada yeni bir *C* olayı tanımlayalım.

C olayı : Çekilen topun siyah olması ise

$A\cap C$ olayı : Çekilen topun hem beyaz hem siyah olması olur. torbada hem siyah hem beyaz top olmadığına göre $A \cap C$ boş kümedir. Yani *A* ve *C* olayları ayrık olaylardır. Dolayısıyla;

$$P(A \cap C) = 0$$

olur.

$A|C$ olayı: Çekilen topun siyah olduğu bilindiğinde beyaz olması.

Bu da boş bir kümedir. Bunun da olasılığın 0 olacağı tabiidir. Fakat bunu şartlı olasılık tanımından hareketle elde edelim.

$$P(A|C) = \frac{P(A \cap C)}{P(C)} = \frac{0}{\frac{4}{11}} = 0$$

Çarpma ve Toplama Kuralları

A ve *B* olayları verildiğinde, iki durum söz konusu olabilir.

$A \cap B$ olayı : *A* ve *B*'nin birlikte olması olayı

$A \cup B$ olayı : *A* veya *B*'nin olması olayı

Burada birinci olayın olasılığı olasılıkların çarpma kuralı ile, ikinci olayın olasılığı ise olasılıkların toplama kuralı ile bulunur;

Çarpma Kuralı $P(A \cap B) = P(A).P(B|A) = P(B).P(A|B)$

Toplama Kuralı $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

olarak yazılır. Eğer *A* ve *B* olayları bağımsız iseler;

$$P(B|A) = P(B)$$

olacağından, çarpa kuralı;

$$P(A \cap B) = P(A).P(B)$$

halini alır. Öte yandan *A* ve *B* olayları ayrık iseler;

$$A \cap B = \varnothing \quad P(A \cap B) = 0$$

olacağından, toplama kuralı da;

$$P(A\cup B)=P(A)+P(B)$$

şeklinde sadeleşir.

Örnek: İçinde 6 beyaz ve 7 siyah top bulunan torbadan iadesiz olarak iki top çekiliyor.

A olayı : Çekilen ilk topun beyaz olması

B olayı : İkinci topun siyah olması

$A\cap B$ olayı : Birinci topun beyaz ikincinin siyah olması

şeklinde tanımlanabilir. *A* ve *B* olayları bağımlı olaylardır. Bu $A\cap B$ olayının olasılığını bağımlı olayların olasılık çarpım kuralından;

$$P(A\cap B)=P(A).P(B|A)=\frac{6}{13}.\frac{7}{12}=\frac{7}{26}=0.27$$

olarak bulunur. Eğer çekilen toplar tekrar torbaya iade ediliyor olsaydı, *A* ve *B* olayları bağımsız olacaklardır. Bu durumda;

$$P(A\cap B)=P(A).P(B)=\frac{7}{13}.\frac{6}{13}=\frac{42}{169}=0.25$$

olarak olasılık bulunurdu.

Örnek: 6 beyaz 7 siyah top bulunan torbadan arka arkaya iadesiz iki top çekilişi deneyinde yine;

A olayı : Çekilen ilk topun beyaz olması

B olayı : İkinci topun siyah olması

olsun. İadesiz seçimden dolayı *A* ve *B* olayları bağımlıdır. Burada;

$A\cup B$ olayı : Çekilen ilk topun beyaz veya ikinci topun siyah olması olarak tanımlanır. Bu olayın olasılığını toplama kuralından;

$$P(A\cup B)=P(A)+P(B)-P(A\cap B)$$

$$=\frac{6}{13}+\frac{7}{13}-\frac{6}{13}.\frac{7}{12}=\frac{114}{156}=0.73$$

buluruz. Topların çekilişi iadeli olsaydı *A* ve *B* olayları bağımsız olaylar olacaktı. Bu durumda $A\cup B$ olayının olasılığı;

$$P(A\cup B)=P(A)+P(B)-P(A\cap B)$$

$$=\frac{6}{13}+\frac{7}{13}-\frac{6}{13}.\frac{7}{13}=\frac{127}{169}=0.75$$

bulunur.

Örnek : Bir zar atma deneyinde *A* olayı çift gelmesi, *B* olayı tek gelmesi olsun. Bu durumda $A\cap B=\varnothing$ olduğundan $P(A\cap B)=0$ ve;

$$P(A\cup B)=P(A)+P(B)=\frac{3}{6}+\frac{3}{6}=1$$

bulunur. Yani zarın atılışında tek veya çift gelmesi kesin bir olaydır. Bir *A* olayı;

$$A\cap B_1,\ A\cap B_2,...,A\cap B_r$$

gibi parçalardan oluşuyor ise;

$$P(A)=P(A\cap B_1)+P(A\cap B_2)+\cdots+P(A\cap B_r)$$

ifadesine *A*'nın toplam olasılığı denir. Burada kesişimler şartlı olasılık ile;

$$P(A)=P(A|B_1)+P(A|B_2)+\cdots+P(A|B_r)$$

yazılabilir. Buradan bir B_j olayı verildiğinde *A*'nın şartlı olasılığı;

$$P(B_j|A)=\frac{P(A|B_j)P(B_j)}{\sum P(A|B_j)P(B_j)}$$

veya

$$P(B_j|A)=\frac{P(A|B_j)P(B_j)}{P(A)}$$

olur ki bu ifade *Bayes Teoremi* olarak bilinir. Teoremi kesişimler cinsinden;

$$P(B_j|A)=\frac{P(A\cap B_j)}{\sum P(A\cap B_j)}=\frac{P(A\cap B_j)}{P(A)}$$

olarak da yazabiliriz.

Örnek: Bir sınıfta 20 kız, 30 erkek öğrenci mevcuttur. Kızların 10'u İngilizce, 7'si Fransızca, 3'ü de Almancayı yabancı dil olarak seçerken erkek öğrencilerin 15'i İngilizceyi, 10’u Fransızcayı, 5’i de Almancayı seçmiştir. Seçilen bir öğrencinin kız olması halinde yabancı dil olarak İngilizceyi seçmiş olması ihtimalini bulalım.

Çözüm:

A olayı	: Seçilen öğrencinin kız olması
B_1 olayı	: Seçilen öğrencinin yabancı dilinin İngilizce olması
B_2 olayı	: Seçilen öğrencinin yabancı dilinin Fransızca olması
B_3 olayı	: Seçilen öğrencinin yabancı dilinin Almanca olması
$B_1 \vert A$ olayı	: Kız olması halinde yabancı dilinin İngilizce olması
$A \vert B_1$ olayı	: Yabancı dilinin İngilizce olması halinde kız olması
$A \vert B_2$ olayı	: Yabancı dilinin Fransızca olması halinde kız olması
$A \vert B_3$ olayı	: Yabancı dilinin Almanca olması halinde kız olması

Bu olayların olasılıkları sırasıyla;

$$P(A) = \frac{20}{50} = 0.40$$

$$P(B_1) = \frac{25}{50} = 0.50$$

$$P(B_2) = \frac{17}{50} = 0.34$$

$$P(B_3) = \frac{8}{50} = 0.16$$

$$P(A|B_1) = \frac{10}{20} = 0.50$$

$$P(A|B_2) = \frac{7}{20} = 0.35$$

$$P(A|B_3) = \frac{3}{20} = 0.15$$

$$P(B_1|A) = ?$$

Bu verilerden $P\left(B_1 \middle| A\right)$ olasılığı:

$$P\left(B_1 \middle| A\right) = \frac{P\left(A \middle| B_1\right) P\left(B_1\right)}{\sum P\left(A \middle| B_j\right) P\left(B_j\right)}$$

$$= \frac{(0.50).(0.50)}{\left[(0.50).(0.50) + (0.35).(0.34) + (0.15).(0.16)\right]}$$

$$= 0.64$$

olarak bulunur.