Olasılık ve İstatistik

HAFTA 6

Özel Kesikli Olasılık Dağılımları:

Binom Dağılımı

Hipergeometrik Dağılım

Poisson Dağılımı

Dr. Öğretim Üyesi Burcu ÇARKLI YAVUZ

bcarkli@sakarya.edu.tr

Binom dağılımı

- Bir denemenin başarılı ve başarısız gibi 2 olası sonucu varsa, bunların olasılıkları sırasıyla p ve 1-p ise bu denemeye ***Bernoulli denemesi*** denir.
- n adet Bernoulli denemesindeki başarılı durum sayısı ***X***, bir ***Binom rastgele değişkeni*** olarak adlandırılır. Bu kesikli rastgele değişkenin olasılık dağılımına ise ***Binom Dağılımı*** denir. Tekrarlanan denemeler birbirinden bağımsızdır.
- p: başarılı olma olasılığını, q=1-p başarısız olma olasılığını göstermek üzere, X binom rastgele değişkeninin olasılık dağılımı şu şekilde hesaplanır:

$$b(X;n,p) = \binom{n}{x} p^x q^{n-x}, x = 0,1,2, \ldots n$$

Örnek

- Belli bir cihazın bir parçasının belirli bir şok karşısında hayatta kalma olasılığı ¾ tür. Test edilen 4 parçanın tam olarak 2 tanesinin hayatta kalma olasılığını bulunuz.

p=3/4 (parçanın hayatta kalma olasılığı)

q=1-3/4=1/4 (parçanın hayatta kalmama olasılığı)

X=2 (tam olarak 2 tanesi hayatta kalacak)

n=4 (toplam 4 parça test edilecek yani 4 deney yapılacak)

$$b\left(2;4,\frac{3}{4}\right)=\binom{4}{2}\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^2=\frac{4!}{2!\,2!}*\frac{3^2}{4^4}=\frac{27}{128}=0.21$$

Örnek

➢Hilesiz bir paranın 12 atılışında 5 yazı, 7 tura gelme olasılığı nedir?

p=1/2 (yazı gelme olasılığı)

q=1-1/2=1/2 (tura gelme olasılığı)

X=5 (yazı sayısı)

n=12 (deney sayısı)

$$b\left(5;12,\frac{1}{2}\right)=\binom{12}{5}\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^7=\frac{12!}{5!\,7!}*\frac{1}{2^5}*\frac{1}{2^7}=0.193$$

Binom Tablosu kullanımı

- Binom tablosunda n (deney sayısı) değeri ve p değeri için tablonun uygun kısmını bul.
- İstenilen X değeri için tablodan değeri oku
- Önceki soruda n=12, p=1/2, x=5 için tablodan ilgili değeri bulalım:

		p										
n	x	.05	.1	.2	.3	.4	.5	.6	.7	.8	.9	.95
12	0	.540	.282	.069	.014	.002						
	1	.341	.377	.206	.071	.017	.003					
	2	.099	.230	.283	.168	.064	.016	.002				
	3	.017	.085	.236	.240	.142	.054	.012	.001			
	4	.002	.021	.133	.231	.213	.121	.042	.008	.001		
	5		.004	.053	.158	.227	.193	.101	.029	.003		
	6			.016	.079	.177	.226	.177	.079	.016		
	7			.003	.029	.101	.193	.227	.158	.053	.004	
	8			.001	.008	.042	.121	.213	.231	.133	.021	.002
	9				.001	.012	.054	.142	.240	.236	.085	.017
	10					.002	.016	.064	.168	.283	.230	.099
	11						.003	.017	.071	.206	.377	.341
	12							.002	.014	.069	.282	.540

Örnek

➢ Bir hastanın az rastlanan bir kan hastalığından kurtulması olasılığı 0.4’tür. 15 kişinin bu hastalığa yakalandığı biliniyor ise

a) Tam olarak 5 kişinin hayatta kalması
b) 3’ten 9’a kadar kişinin hayatta kalması (9 dahil değil)
c) En az 10 kişinin hayatta kalması olasılığını bulunuz.

Çözüm

Hayatta kalan kişi sayısı X olsun:

n=15, p=0.4

a) P(X=5)=0.186

b) P(3≤X≤8)=P(X=3)+P(X=4)+…+P(X=8)

=0.063+0.127+0.186+0.207+0.177+0.118=0.878

c) P(X≥10)=0.024+0.007+0.002+0+0+0=0.033 (binom tab. x=10,11,…,15)

yada 1-P(X<10)

=1-(0+0.005+0.022+0.063+0.127+0.186+0.207+0.177+0.118+0.061)
=1-0,966=0.034

n=15 için Binom tablosu

n	x	p										
		.05	.1	.2	.3	.4	.5	.6	.7	.8	.9	.95
15	0	.463	.206	.035	.005							
	1	.366	.343	.132	.031	.005						
	2	.135	.267	.231	.092	.022	.003					
	3	.031	.129	.250	.170	.063	.014	.002				
	4	.005	.043	.188	.219	.127	.042	.007	.001			
	5	.001	.010	.103	.206	.186	.092	.024	.003			
	6		.002	.043	.147	.207	.153	.061	.012	.001		
	7			.014	.081	.177	.196	.118	.035	.003		
	8			.003	.035	.118	.196	.177	.081	.014		
	9			.001	.012	.061	.153	.207	.147	.043	.002	
	10				.003	.024	.092	.186	.206	.103	.010	.001
	11				.001	.007	.042	.127	.219	.188	.043	.005
	12					.002	.014	.063	.170	.250	.129	.031
	13						.003	.022	.092	.231	.267	.135
	14							.005	.031	.132	.343	.366
	15								.005	.035	.206	.463

Binom Dağılımının Ortalaması ve Varyansı

- b(X;n,p) binom dağılımının, ortalaması (beklenen değeri):

$$\mu = n * p$$

- Varyansı:

$$\sigma^2 = n * p * q$$

şeklinde hesaplanır.

Örnek

- Kan hastalığı örneğinde verilen binom rastgele değişkeninin ortalamasını ve varyansını bulunuz.

n=15

p=0.4

q=0.6

$$\mu = n * p = 15 * 0.4 = 6$$

$$\sigma^2 = n * p * q = 15 * 0.4 * 0.6 = 3.6$$

Hipergeometrik Dağılım

- Binom dağılımında tekrarlanan deneyler birbirinden bağımsızdı ve p her deneme için aynı değere sahipti.
- Hipergeometrik dağılımda iadesiz örnekleme yapılmaktadır.
- n çaplı rastgele örneklem, N nesne arasından iadesiz olarak seçilir. N tane nesnede k adet başarı, N-k adet başarısızlık olarak sınıflandırılabilir.
- Genellikle N nesneden n çaplı rastgele bir örnek seçildiğinde, başarılı olan k nesneden x tanesinin, başarısız olan N-k nesneden n-x tanesinin seçilmesi olasılığı hipergeometrik olasılık kavramı ile açıklanabilir.
- X hipergeometrik rastgele değişkeninin olasılık dağılımı şu şekilde bulunur:

$$h(x; N, n, k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}, \quad \max\{0, n-(N-k)\} \leq x \leq \min\{n, k\}.$$

Örnek

- Bir denetmen, bir hava kirliliği araştırmasının bir parçası olarak bir şirketin 24 kamyonundan 6'sının egzoz gazı çıkarımını inceleyecektir. Eğer kamyonlardan 4'ü aşırı hava kirliliği yaratıyorsa, bunlardan hiçbirinin denetmenin örneklemine girmemesi olasılığı nedir?

N=24 (toplam kamyon sayısı)

n=6 (örneklem boyutu)

k=4 (hava kirliliğine sebep olan kamyon sayısı)(N-k=24-4=20 uygun)

X=0 (örneklemde hava kirliliğine yol açan kamyon bulunmayacak)

$$h(0;24,6,4)=\frac{\binom{4}{0}\binom{20}{6}}{\binom{24}{6}}=0.2880$$

Not: Hipergeometrik dağılımlar tablo ile çözülmüyor.

Örnek

- 40 parça içeren paketlerin her biri 3 ya da daha fazla kusurlu parça içeriyor ise, kabul edilemez sayılmaktadır.
- Bir pakete ilişkin örnekleme prosedürü, rastgele 5 parça seçilmesi ve bu 5 parçanın içinde kusurlu bulunursa kutunun reddedilmesi biçimindedir. Bir paketin içinde 3 kusurlu parça varsa, seçilen örneklem içinde tam olarak 1 kusurlu parça bulunması olasılığı nedir?

N=40 (toplam parça sayısı)

n=5 (örneklem boyutu)

k=3 kusurlu parça sayısı (N-k = 40-3=37 kusurlu olmayan parça sayısı)

X=1

$$h(1;40,5,3)=\frac{\binom{3}{1}\binom{37}{4}}{\binom{40}{5}}=0.3011$$

- Bu şekilde ürün kontrolü yapılırsa, kusurlu olması gereken bir paket %30 olasılıkla kusurlu olarak işaretlenecektir. Ürün kontrolü için daha uygun bir yol seçilebilir.

Hipergeometrik dağılımın ortalaması ve varyansı

- h(X;N,n,k) hipergeometrik dağılımın ortalaması ve varyansı aşağıdaki gibi tanımlanır:

$$\mu = \frac{n * k}{N}$$

$$\sigma^2 = \frac{N-n}{N-1} * n * \frac{k}{N} * (1 - \frac{k}{N})$$

Poisson Dağılımı

- Belli bir zaman aralığında yada belirli bir bölgede ortaya çıkan sonuçların sayısı olan bir X rastgele değişkeninin sayısal değerlerinin elde edildiği deneyler Poisson deneyleri olarak adlandırılır.
- Bir Poisson deneyi süresince ortaya çıkan sonuçların sayısı X, bir ***Poisson rastgele değişkeni*** ve X'in olasılık dağılımı ***Poisson dağılımı*** olarak adlandırılır. Poisson dağılımı şu şekilde hesaplanır:

$$p(X;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}, x = 0,1,2,\ldots$$

Burada λ ortalama sonuç sayısıdır. e=2.71'dir.

Örnek

- Bir laboratuvar deneyinde 1 milisaniye içinde bir sayaç üzerinden geçen radyoaktif parçacıkların ortalama sayısı 4'tür. Belli bir milisaniyede 6 parçacığın sayaca girmesi olasılığı nedir?

λ=4 (ortalama)

X=6 (istenilen)

$$p(6;4) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{4^6 e^{-4}}{6!} = 0.1042$$

Poisson Tablosu ile çözüm

	λ									
X	**3.1**	**3.2**	**3.3**	**3.4**	**3.5**	**3.6**	**3.7**	**3.8**	**3.9**	**4.0**
0	.0450	.0408	.0369	.0334	.0302	.0273	.0247	.0224	.0202	.0183
1	.1397	.1304	.1217	.1135	.1057	.0984	.0915	.0850	.0789	.0733
2	.2165	.2087	.2008	.1929	.1850	.1771	.1692	.1615	.1539	.1465
3	.2237	.2226	.2209	.2186	.2158	.2125	.2087	.2046	.2001	.1954
4	.1734	.1781	.1823	.1858	.1888	.1912	.1931	.1944	.1951	.1954
5	.1075	.1140	.1203	.1264	.1322	.1377	.1429	.1477	.1522	.1563
6	.0555	.0608	.0662	.0716	.0771	.0826	.0881	.0936	.0989	.1042
7	.0246	.0278	.0312	.0348	.0385	.0425	.0466	.0508	.0551	.0595
8	.0095	.0111	.0129	.0148	.0169	.0191	.0215	.0241	.0269	.0298
9	.0033	.0040	.0047	.0056	.0066	.0076	.0089	.0102	.0116	.0132
10	.0010	.0013	.0016	.0019	.0023	.0028	.0033	.0039	.0045	.0053
11	.0003	.0004	.0005	.0006	.0007	.0009	.0011	.0013	.0016	.0019
12	.0001	.0001	.0001	.0002	.0002	.0003	.0003	.0004	.0005	.0006
13	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0002	.0002
14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

Örnek

- Bir hastanenin acil servisine 10 dakikalık bir zaman aralığında ortalama 3 hasta gelmektedir. Bu zaman aralığında,

a) Hasta gelmemesi olasılığı nedir? (P(X=0)=?)

b) 2 hasta gelmesi olasılığı nedir? (P(X=2)=?)

c) En fazla 5 hasta gelmesi olasılığı nedir? (P(X≤5)=?)

$$p(0;3) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{3^0 e^{-3}}{0!} = 0.05$$

$$p(2;3) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{3^2 e^{-3}}{2!} = 0.224$$

$$p(x \le 5) = p(0;3) + p(1;3) + p(2;3) + p(3;3) + p(4;3) + p(5;3)$$

= 0.0498+0.1494+0.2240+0.2240+0.1680+0.1008=0.916

λ=3 için Poisson tablosu

	λ									
X	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
0	.1225	.1108	.1003	.0907	.0821	.0743	.0672	.0608	.0550	.0498
1	.2572	.2438	.2306	.2177	.2052	.1931	.1815	.1703	.1596	.1494
2	.2700	.2681	.2652	.2613	.2565	.2510	.2450	.2384	.2314	.2240
3	.1890	.1966	.2033	.2090	.2138	.2176	.2205	.2225	.2237	.2240
4	.0992	.1082	.1169	.1254	.1336	.1414	.1488	.1557	.1622	.1680
5	.0417	.0476	.0538	.0602	.0668	.0735	.0804	.0872	.0940	.1008
6	.0146	.0174	.0206	.0241	.0278	.0319	.0362	.0407	.0455	.0504
7	.0044	.0055	.0068	.0083	.0099	.0118	.0139	.0163	.0188	.0216
8	.0011	.0015	.0019	.0025	.0031	.0038	.0047	.0057	.0068	.0081
9	.0003	.0004	.0005	.0007	.0009	.0011	.0014	.0018	.0022	.0027
10	.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0005	.0006	.0008
11	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0002
12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

Örnek

- Belirli bir limana her gün ortalama on petrol tankeri gelmektedir. Limandaki tesisler günde en fazla 15 tankerin işlemini yapabilmektedir. Belirli bir günde tankerlerin geri çevrilmesi olasılığı nedir?

$\lambda=10$

$P(X>15)=1-P(X\leq 15)=?$

$P(X>15)=0.0217+0.0128+0.0071+0.0037+0.0019+0.0009+0.0004+0.0002+0.0001=0.0488$

yada

$1-P(X\leq 15)=1-(0+0.0005+0.0023+0.0076+0.0189+0.0378+0.0631+0.0901+0.1126$

$+0.1251+0.1251+0.1137+0.0948+0.0729+0.0521+0.0347)=0.0487$

λ=10 için Poisson tablosu

	λ									
x	9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10
0	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0000
1	.0010	.0009	.0009	.0008	.0007	.0007	.0006	.0005	.0005	.0005
2	.0046	.0043	.0040	.0037	.0034	.0031	.0029	.0027	.0025	.0023
3	.0140	.0131	.0123	.0115	.0107	.0100	.0093	.0087	.0081	.0076
4	.0319	.0302	.0285	.0269	.0254	.0240	.0226	.0213	.0201	.0189
5	.0581	.0555	.0530	.0506	.0483	.0460	.0439	.0418	.0398	.0378
6	.0881	.0851	.0822	.0793	.0764	.0736	.0709	.0682	.0656	.0631
7	.1145	.1118	.1091	.1064	.1037	.1010	.0982	.0955	.0928	.0901
8	.1302	.1286	.1269	.1251	.1232	.1212	.1191	.1170	.1148	.1126
9	.1317	.1315	.1311	.1306	.1300	.1293	.1284	.1274	.1263	.1251
10	.1198	.1210	.1219	.1228	.1235	.1241	.1245	.1249	.1250	.1251
11	.0991	.1012	.1031	.1049	.1067	.1083	.1098	.1112	.1125	.1137
12	.0752	.0776	.0799	.0822	.0844	.0866	.0888	.0908	.0928	.0948
13	.0526	.0549	.0572	.0594	.0617	.0640	.0662	.0685	.0707	.0729
14	.0342	.0361	.0380	.0399	.0419	.0439	.0459	.0479	.0500	.0521
15	.0208	.0221	.0235	.0250	.0265	.0281	.0297	.0313	.0330	.0347
16	.0118	.0127	.0137	.0147	.0157	.0168	.0180	.0192	.0204	.0217
17	.0063	.0069	.0075	.0081	.0088	.0095	.0103	.0111	.0119	.0128
18	.0032	.0035	.0039	.0042	.0046	.0051	.0055	.0060	.0065	.0071
19	.0015	.0017	.0019	.0021	.0023	.0026	.0028	.0031	.0034	.0037
20	.0007	.0008	.0009	.0010	.0011	.0012	.0014	.0015	.0017	.0019
21	.0003	.0003	.0004	.0004	.0005	.0006	.0006	.0007	.0008	.0009
22	.0001	.0001	.0002	.0002	.0002	.0002	.0003	.0003	.0004	.0004
23	.0000	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0002	.0002
24	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001

Örnek

- Bir ciltçinin ciltlediği kitapların yüzde 2’sinin cildi bozuktur. Bu ciltçinin ciltlediği 400 kitaptan 5 tanesinin bozuk çıkma olasılığı nedir?

λ=400*2/100=8

X=5

$$p(5;8) = \frac{8^5 e^{-8}}{5!} = 0.092$$

Yada Poisson tablosundan sonuç okunabilir: 0.0916

λ=8 için Poisson tablosu

	λ									
X	7.1	7.2	7.3	7.4	7.5	7.6	7.7	7.8	7.9	8.0
0	.0008	.0007	.0007	.0006	.0006	.0005	.0005	.0004	.0004	.0003
1	.0059	.0054	.0049	.0045	.0041	.0038	.0035	.0032	.0029	.0027
2	.0208	.0194	.0180	.0167	.0156	.0145	.0134	.0125	.0116	.0107
3	.0492	.0464	.0438	.0413	.0389	.0366	.0345	.0324	.0305	.0286
4	.0874	.0836	.0799	.0764	.0729	.0696	.0663	.0632	.0602	.0573
5	.1241	.1204	.1167	.1130	.1094	.1057	.1021	.0986	.0951	.0916
6	.1468	.1445	.1420	.1394	.1367	.1339	.1311	.1282	.1252	.1221
7	.1489	.1486	.1481	.1474	.1465	.1454	.1442	.1428	.1413	.1396
8	.1321	.1337	.1351	.1363	.1373	.1382	.1388	.1392	.1395	.1396
9	.1042	.1070	.1096	.1121	.1144	.1167	.1187	.1207	.1224	.1241
10	.0740	.0770	.0800	.0829	.0858	.0887	.0914	.0941	.0967	.0993
11	.0478	.0504	.0531	.0558	.0585	.0613	.0640	.0667	.0695	.0722
12	.0283	.0303	.0323	.0344	.0366	.0388	.0411	.0434	.0457	.0481
13	.0154	.0168	.0181	.0196	.0211	.0227	.0243	.0260	.0278	.0296
14	.0078	.0086	.0095	.0104	.0113	.0123	.0134	.0145	.0157	.0169
15	.0037	.0041	.0046	.0051	.0057	.0062	.0069	.0075	.0083	.0090

Örnek

- Belli bir metal levhada 10m2 başına ortalama 5 kusur vardır. 15m2 levhada en az 6 kusur olma olasılığı kaçtır?

$\lambda=5$ (10m2 için)

$\lambda=7.5$ (15m2 için)

$P(X\geq6)=1-P(X\leq5)$

Poisson tablosuna baktığımızda $1-P(X\leq5)$ hesabı daha kolay olacaktır:

$1-(0.0006+0.0041+0.0156+0.0389+0.0729+0.1094)=0.7585$

λ=7.5 için Poisson tablosu

	λ									
x	7.1	7.2	7.3	7.4	7.5	7.6	7.7	7.8	7.9	8.0
0	.0008	.0007	.0007	.0006	.0006	.0005	.0005	.0004	.0004	.0003
1	.0059	.0054	.0049	.0045	.0041	.0038	.0035	.0032	.0029	.0027
2	.0208	.0194	.0180	.0167	.0156	.0145	.0134	.0125	.0116	.0107
3	.0492	.0464	.0438	.0413	.0389	.0366	.0345	.0324	.0305	.0286
4	.0874	.0836	.0799	.0764	.0729	.0696	.0663	.0632	.0602	.0573
5	.1241	.1204	.1167	.1130	.1094	.1057	.1021	.0986	.0951	.0916
6	.1468	.1445	.1420	.1394	.1367	.1339	.1311	.1282	.1252	.1221
7	.1489	.1486	.1481	.1474	.1465	.1454	.1442	.1428	.1413	.1396
8	.1321	.1337	.1351	.1363	.1373	.1382	.1388	.1392	.1395	.1396
9	.1042	.1070	.1096	.1121	.1144	.1167	.1187	.1207	.1224	.1241
10	.0740	.0770	.0800	.0829	.0858	.0887	.0914	.0941	.0967	.0993
11	.0478	.0504	.0531	.0558	.0585	.0613	.0640	.0667	.0695	.0722
12	.0283	.0303	.0323	.0344	.0366	.0388	.0411	.0434	.0457	.0481
13	.0154	.0168	.0181	.0196	.0211	.0227	.0243	.0260	.0278	.0296
14	.0078	.0086	.0095	.0104	.0113	.0123	.0134	.0145	.0157	.0169
15	.0037	.0041	.0046	.0051	.0057	.0062	.0069	.0075	.0083	.0090
16	.0016	.0019	.0021	.0024	.0026	.0030	.0033	.0037	.0041	.0045
17	.0007	.0008	.0009	.0010	.0012	.0013	.0015	.0017	.0019	.0021
18	.0003	.0003	.0004	.0004	.0005	.0006	.0006	.0007	.0008	.0009
19	.0001	.0001	.0001	.0002	.0002	.0002	.0003	.0003	.0003	.0004
20	.0000	.0000	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0002
21	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001

Kaynaklar

- Doç. Dr. Melda Akın, Matematiksel İstatistik Ders Notu, İstanbul Üniversitesi
- Mühendisler ve Fen Bilimciler için Olasılık ve İstatistik, Ronald E. Walpole, Çeviri Editörü: Prof. Dr. M. Akif BAKIR, 9. baskıdan çeviri, Palme Yayıncılık, 2016.
- İstatistiksel Formüller ve Tablolar, Başkent Üniversitesi İktisadi ve İdari Bilimler Fakültesi, 2005