ÖZEL DAĞILIMLAR

5.1. Giriş

Bir araştırmacı için rastlantı değişkeninin olasılık fonksiyonu ya da olasılık yoğunluk fonksiyonları önemli kavramlardır. Araştırmacının yaptığı deney ya da gözlemler için bir olasılık fonksiyonu yaratmaya çalışması, hem araştırmada önemli bazı detayların gözden kaçmasına hem de daha önce geliştirilmiş bir olasılık fonksiyonunu elde etmek için zaman ve kaynak israfına neden olacaktır.

Bu amaçla olasılık kuramında belirli özelliklere bağlı olarak kuramsal dağılımlar geliştirilmiştir. Kuramsal dağılımlar; kesikli dağılımlar ve sürekli dağılımlar olmak üzere iki ana başlık altında toplanmaktadır.

Bu bölümün amacı, ele alınan dağılımın hangi durumlarda kullanıldığını ortaya koymak, parametrelerini, olasılık fonksiyonunu, moment çıkaran fonksiyonunu incelemek, beklenen değer ve varyansının neye eşit olduğunu ve nasıl bulunduğuna cevap. vermektir.

5.2. Kesikli Dağılımlar

Bu bölümde kesikli kuramsal dağılımlar incelenecektir. Kesikli kuramsal dağılımlar şunlardır.

- Kesikli Uniform (Tek düze) dağılım
- Bernoulli dağılımı
- Binom (İki terimli) dağılım
- Geometrik dağılım
- Negatif Binom (Eksi İki Terimli) dağılım
- Hipergeometrik dağılım
- Poisson dağılımı

5.2.1. Kesikli Düzgün (Uniform) Dağılım

Bir rastlantı değişkeninin eşit olasılıkla k farklı değer alıyorsa, bu rastlantı değişkeninin kesikli uniform dağılıma uyduğu söylenir. Kesikli Uniform dağılımın olasılık fonksiyonu $P_X(x)=\frac{1}{k};\quad x=1,2,.....,k$ için biçimindedir. Burada k, dağılımın parametresidir. Dağılım ortalaması, varyansı ve Moment Çıkaran Fonksiyonu (M.Ç.F) aşağıdaki gibidir.

$$E(X)=\sum_{x=1}^{k} x\left(\frac{1}{k}\right)=\left(\frac{1}{k}\right)\sum_{x=1}^{k} x=\frac{k(k+1)}{2k}=\frac{(k+1)}{2}$$ olur. Benzeri şekilde

$$E(X^2)=\sum_{x=1}^{k} x^2\left(\frac{1}{k}\right)=\left(\frac{1}{k}\right)\sum_{x=1}^{k} x^2=\frac{k(k+1)(2k+1)}{6k}=\frac{(k+1)(2k+1)}{6}$$

$$Var(X)=E(X^2)-\left[E(X)\right]^2=\frac{(k+1)(2k+1)}{6}-\left(\frac{k+1}{2}\right)^2=\frac{k^2-1}{12},$$

$$M_X(t)=E\left(e^{tX}\right)=\sum_{x=1}^{k}\frac{1}{k}e^{tx}=\frac{1}{k}\left(e^t+e^{2t}+e^{3t}+...+e^{kt}\right)$$

$$=\frac{1}{k}\left(e^t\right)\left(1+e^t+(e^t)^2+(e^t)^3+...+(e^t)^{(k-1)}\right)=\frac{e^t\left(1-e^{kt}\right)}{k(1-e^t)}$$

Böyle bir dağılım, örneğin bir zar atma denemesinde üste gelecek yüzdeki noktaların sayıları için uygun bir dağılımdır.

5.2.2. Bernoulli Dağılımı

Bernoulli dağılımı bir denemenin iki sonucu (iyi–kötü, olumlu-olumsuz) olması durumunda kullanılan bir dağılımdır. İki sonuçlu bir Bernoulli deneyinde ilgilenilen sonuca "başarı", diğer sonuca, "başarısızlık" dersek bunu X değişkeni olarak aşağıdaki gibi ifade edebiliriz.

Bir deneyin başarılı olması olasılığı p ile gösterilirse X değişkeninin olasılık fonksiyonu

$$P_X(x;p) = P_x(x) = \begin{cases} p^x(1-p)^{1-x}; & x = 0,1 \\ 0; & d.d \end{cases}$$

biçiminde verilir. Bernoulli dağılımın parametresi p'dir. Dağılımın ortalaması, varyansı ve M.Ç.F'u aşağıdaki gibidir.

$$E(X) = p = \mu\,,\ Var(X) = \sigma^2 = p(1-p) = pq$$

$$M_x(t) = pe^t + (1-p) = pe^t + q\,,\ \text{burada}\ q = 1-p\,.$$

Bernoulli dağılımında deney bir kez yapılıp başarı sonucu ile ilgilenilmektedir. Eğer deney n kez ve birbirinden bağımsız olmak üzere tekrarlanırsa ve yine başarı sonucu ile ilgilenilirse bu durumda Bernoulli dağılımının özel bir durumu olan Binom dağılımının konu edilmesi gerekir.

5.2.3. Binom (İki Terimli) Dağılım

Binom dağılımının kullanım alanı oldukça geniştir. Binom dağılımdan yararlanabilmek için bazı koşulların sağlanması gerekir.

- Rastgele deney aynı koşullar altında n kez tekrarlanmalıdır,
- Her deneyin başarı-başarısızlık gibi iki sonucu olmalıdır,
- Bir deneyde ilgilenilen sonucu (başarı) elde etme olasılığı p, diğer sonucu (başarısız) elde etme olasılığı 1-p=q, sabit olmalıdır,
- Deneme sonuçları birbirinden bağımsız olmalıdır.

Bu dört koşulun sağlandığı n tane Bernoulli deneyinde X değişkenin alacağı değerler: 0,1,.....,n olabilir.

Bu şartları sağlayan n Bağımsız Bernoulli deneyi sonucunda elde edilecek toplam başarı sayısı olan X 'in olasılık fonksiyonu aşağıdaki şekilde verilir.

$$P(x;n,p)=\begin{cases}\binom{n}{x}p^x q^{(n-x)}; & x=0,1,2,3,\ldots n\\ 0; & d.d\end{cases}$$

Burada , $\binom{n}{x}=\dfrac{n!}{x!(n-x)!}$ olarak alınmalıdır.

Binom dağılımın ortalaması, varyansı ve moment çıkaran fonksiyonu aşağıda verildiği gibidir.

$$\mu=E(X)=np\,,\ Var(X)=\sigma^2=npq \ \text{ ve } \ M_X(t)=\left[pe^t+q\right]^n$$

İspat:

$$\mu = E(X) = \sum_{x=0}^{n} x \frac{n!}{x!(n-x)!} p^x q^{(n-x)} = \sum_{x=1}^{n} x \frac{n!}{x!(n-x)!} p^x q^{(n-x)}$$

$$= \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!} p^x q^{(n-x)}$$

$x - 1 = t$ olarak alınırsa,

$$\mu = E(X) = \sum_{t=0}^{n-1} \frac{n!}{(t)!(n-1-t)!} p^{t+1} q^{(n-1-t)} = np \sum_{t=0}^{n-1} \frac{(n-1)!}{(t)!(n-1-t)!} p^t q^{(n-1-t)}$$

$$= np\left[p+q\right]^{(n-1)} = np$$

$$E(X^2) = \sum_{x=0}^{n} x^2 \frac{n!}{x!(n-x)!} p^x q^{(n-x)} = \sum_{x=1}^{n} x^2 \frac{n!}{x!(n-x)!} p^x q^{(n-x)}$$

$$= \sum_{x=1}^{n} x \frac{n!}{(x-1)!(n-x)!} p^x q^{(n-x)}$$

$x - 1 = t$ olarak alınırsa:

$$E(X^2) = \sum_{x=1}^{n} x \frac{n!}{(x-1)!(n-x)!} p^x q^{(n-x)} = \sum_{t=0}^{n-1} (t+1) \frac{n!}{(t)!(n-1-t)!} p^{t+1} q^{(n-1-t)}$$

$$= \sum_{t=0}^{n-1} (t) \frac{n!}{(t)!(n-1-t)!} p^{t+1} q^{(n-1-t)} + \sum_{t=0}^{n-1} (1) \frac{n!}{(t)!(n-1-t)!} p^{t+1} q^{(n-1-t)}$$

$$= np \sum_{t=0}^{n-1} (t) \frac{(n-1)!}{(t)!(n-1-t)!} p^t q^{(n-1-t)} + np \sum_{t=0}^{n-1} (1) \frac{(n-1)!}{(t)!(n-1-t)!} p^t q^{(n-1-t)}$$

$$= np[(n-1)p] + np = n^2 p^2 - np^2 + np$$

Buradan

$\sigma^2 = E(X^2) - [E(X)]^2 = n^2 p^2 - np^2 + np - (np)^2 = np(1-p) = npq$ olur.

$$M_X(t) = E(e^{tX}) = \mu = E(X) = \sum_{x=0}^{n} e^{tx} \frac{n!}{x!(n-x)!} p^x q^{(n-x)}$$

$$= \sum_{x=0}^{n} \frac{n!}{x!(n-x)!} (pe^t)^x q^{(n-x)} = [pe^t + q]^n$$

olur.

Örnek 5.1: Bir para atma denemesinde, para yazı gelirse 0 ve tura gelirse 1 ile gösterilsin.

Bu durumda X değişkenin olasılık fonksiyonunu yazınız.

$$X = \begin{cases} 0, & yazı\,ise \\ 1, & tura\,ise \end{cases}$$

Olası iki durum var, Bernoulli dağılımına uyar.

$$P_X(x, 1/2) = \begin{cases} 1/2, & x = 0\,için \\ 1/2, & x = 1\,için \\ 0, & d.d. \end{cases}$$

Örnek 5.2: Rastgele seçilmiş üç çocuklu bir ailedeki çocukların ailedeki kız çocuklarının sayısı X olsun. X alabileceği değerlerin 0,1,2,3 olacağı kolayca söylenebilir.

Herhangi bir üç çocuklu aile için aşağıdaki olasılıkları hesaplayalım.

a) $P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{3}{0} p^0 q^3 = 1 - q^3$ Eğer $p = q = \frac{1}{2}$ alınırsa

$P(X \geq 1) = 1 - P(X = 0) = 1 - q^3 = 1 - \left(\frac{1}{2}\right)^3 = \frac{7}{8}$ olur.

b) $P(X = 1) = \binom{3}{1} p^1 q^2 = 3pq^2$ olur. Burada da $p = q = \frac{1}{2}$ alınırsa

$P(X = 1) = 3pq^2 = 3\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^2 = \frac{3}{8}$ olur.

c) Üç çocuklu aileler içinden rastgele seçilen 1000 ailedeki kız çocuklarının sayıları saptanacak olsa bu 1000 aileden kaç tanesinde, ortalama olarak, tamı tamına birer kız çocuğu vardır?

T= Üç çocuklu 1000 aile içinde 1 kız (ve 2 erkek) çocuklu aileler sayısı olsun.

Bu şekilde tanımlanan T değişkeninin dağılımı da Binom olur. Ancak burada T'nin parametreleri $n_T = 1000$ ve $p_T = P(X = 1) = \frac{3}{8}$ olur. Yani,

$T \sim Binom(n_T = 1000, p_T = \frac{3}{8})$

$E(T) = n_T p_T = 1000\left(\frac{3}{8}\right) = 375$ aile olarak bulunur

Örnek 5.3: Bir atıcının hedefi vurma olasılığı $\frac{2}{3}$'tür. X hedefi vurma sayısını göstermek üzere ve atışların bağımsız olduğu varsayımı altında 8 defa atış yapıldığında,

a) Hedefi tam üç defa vurma,

b) Hedefi en az iki, en çok altı defa vurma,

c) Hedefi üç defadan fazla vurma olasılıklarını bulunuz.

Cevaplar:

a) $P(X=3)=\binom{8}{3}\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^5=\frac{448}{6561}\cong 0.068$

b) $P(2\le X\le 6)=P(X=2)+.....+P(X=6)$

$$=\sum_{X=2}^{6}\binom{8}{x}\left(\frac{2}{3}\right)^x\left(\frac{1}{3}\right)^{8-x}\cong 0,802$$

c) $P(X>3)=P(X\ge 4)=\sum_{x=4}^{8}\binom{8}{x}\left(\frac{2}{3}\right)^x\left(\frac{1}{3}\right)^{8-x}\cong 0,912$

Örnek 5.4: X değişkenin M.Ç.F'u $M_X(t)=\left(\frac{1}{3}e^t+\frac{2}{3}\right)^4$ olmak üzere P(X<2) olasılığını ve E(X), ve Var(X)'i bulunuz.

Binom dağılımının M.Ç.F'u $M_X(t)=\left(pe^t+p\right)^n$ olduğuna göre $p=\frac{1}{3}, q=\frac{2}{3}, n=4$ olur. Buna göre,

$$P(X<2)=P(X=0)+P(X=1)=\sum_{x=0}^{1}\binom{4}{x}\left(\frac{1}{3}\right)^x\left(\frac{2}{3}\right)^{4-x}$$

$$=\binom{4}{0}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^4+\binom{4}{1}\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^3=\frac{16}{81}+\frac{32}{81}=\frac{48}{81}\cong 0,5926,\ bulunur$$

$$E(X)=np=4\left(\frac{1}{3}\right)=\frac{4}{3},\quad Var(X)=npq=4.\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)=\frac{8}{9}\quad bulunur.$$

5.2.4. Geometrik Dağılım

Geometrik dağılımda rastsal değişken, bağımsız Bernoulli deneylerinde, ilk başarı elde edilinceye kadar gerekli deney sayısı, X, dir. Bu değişkenin alacağı değerler $x = 1,2,3,.....$ olacaktır.

Art arda birbirinden bağımsız olarak ve aynı koşullarda tekrarlanan Bernoulli deneylerinde istenen soncun (başarının) ilk defa elde edilmesi için yapılan deney sayısına X dersek, buna geometrik dağılıma sahip değişken denir.

Binom dağılımında deney sayısı sabit ve istenen sonucun(başarının) sayısı bir değişken iken; geometrik dağılımda deney sayısı değişken, fakat istenen sonuç (başarı sayısı) bir(1)'dir.

Örneğin, bir para atma denemesinde, yazı tarafı ilk defa gelinceye kadar yapılan denemelerin sayısı bir geometrik değişkendir. Benzeri şekilde, hedefe atış yapan bir atıcının hedefi ilk kez vurması için yaptığı atış sayısı da geometrik değişken olarak algılanabilir.

İlk (x-1) denemede istenen sonuca ulaşılamaması (başarısızlık) ve x'inci denemede istenen sonuca ulaşılması (başarı) durumunda geometrik dağılımın olasılık fonksiyonu,

$$P_x(x;p) = P_x(x) = \begin{cases} (1-p)^{(x-1)}p & x = 1,2,3,4,.... \\ 0, & dd \end{cases}$$

biçiminde tanımlanır. Dağılım parametresi p'dir. Ortalaması, varyansı ve M.Ç.Fu aşağıda verildiği gibidir.

$$E(X) = \frac{1}{p}, \quad Var(X) = \sigma^2 = \frac{1-p}{p^2} \text{ ve } M_X(t) = \frac{pe^t}{1-qe^t} .$$

İspat:

$$E(X)=\sum_{x=1}^{\infty}xq^{(x-1)}p=p\sum_{x=1}^{\infty}\frac{d}{dq}q^{x}=p\left(\frac{d}{dq}\sum_{x=1}^{\infty}q^{x}\right)=p\frac{d}{dq}\left(q\frac{1}{1-q}\right)$$

$$=p\frac{(1-q)+q}{(1-q)^2}=\frac{p}{(1-q)^2}=\frac{1}{p}$$

$E(X^2)=\sum_{x=1}^{\infty}x^2q^{(x-1)}p$ ve burada $(x-1)=t$ alınırsa

$$E(X^2)=\sum_{t=0}^{\infty}(t+1)^2q^tp=\sum_{t=0}^{\infty}(t^2+2t+1)q^tp=\sum_{t=0}^{\infty}t^2q^tp+2\sum_{t=0}^{\infty}tq^tp+\sum_{t=0}^{\infty}(1)q^tp$$

$$=\sum_{t=1}^{\infty}t^2q^tp+2\sum_{t=1}^{\infty}tq^tp+\sum_{t=0}^{\infty}(1)q^tp=q\sum_{t=1}^{\infty}t^2q^{t-1}p+2q\sum_{t=1}^{\infty}tq^{t-1}p+\sum_{t=0}^{\infty}(1)q^tp$$

$$=qE(X^2)+2qE(X)+1$$

$$E(X^2)=qE(X^2)+2q\left(\frac{1}{p}\right)+1\Rightarrow(1-q)E(X^2)=\frac{2q}{p}+1=\frac{2q+p}{p}=\frac{(1+q)}{p}$$

$$\Rightarrow E(X^2)=\frac{(1+q)}{p^2}$$

$$Var(X)=\sigma^2=\frac{(1+q)}{p^2}-\left(\frac{1}{p}\right)^2=\frac{q}{p^2}$$

$$M_X(t)=\sum_{x=1}^{\infty}e^{tx}q^{(x-1)}p=\frac{p}{q}\sum_{x=1}^{\infty}(qe^t)^x=\frac{p}{q}(qe^t)\sum_{x=0}^{\infty}(qe^t)^x=pe^t\left(\frac{1}{1-qe^t}\right)$$

Örnek 5.5: Bir zar 3 noktalı yüzü üste gelinceye kadar, X, atılacaktır.

a) X'in olasılık fonksiyonu bulunuz.

b) X'in ortalamasını ve standart sapmasını bulunuz.

c) 3 noktalı yüzün ilk defa 4.cü denemede üste gelme olasılığını bulunuz.

P (3 noktalı yüzün üste gelmesi) $=\frac{1}{6}=$ p, ve q$=\frac{5}{6}$'dır.

Cevaplar:

a) $$P_X(x)=\begin{cases}\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^{x-1}, & x=1,2,3,\ldots \\ 0, & dd\end{cases}$$

b) $$E(X)=\frac{1}{p}=\frac{1}{\frac{1}{6}}=6 \qquad Var(X)=\sigma^2=\frac{1-p}{p^2}=\frac{\frac{5}{6}}{\frac{1}{36}}=30 \text{ ve } \sigma=\sqrt{30}\cong 5.477$$

c) $$P(X=4)=\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^{4-1}=\frac{125}{(6)^4}=\frac{125}{1296}\cong 0,0965 \text{ bulunur.}$$

5.2.5. Negatif Binom Dağılım

Bağımsız Bernoulli deneylerinde eğer ilk başarı değil de, k tane başarı elde etme söz konusu ise geometrik dağılımın genelleştirilmiş hali olan genelleştirilmiş binom, (negatif binom) dağılımından yararlanılır. k'nıncı istenen sonuç (k nıncı başarı) elde edilinceye kadar tekrar edilmesi gereken deneme sayısına X dersek, buna negatif Binom değişkeni denir.

Bir deneyde istenen sonucun ortaya çıkması olasılığı p, istenen sonucun ortaya çıkmaması olasılığı q = (1-p) ve istenen sonucun elde edilme sayısına da k dersek X rastlantı değişkenini olasılık fonksiyonu:

$$P_X(x;k,p) = P_X(x) = \begin{cases} \binom{x-1}{k-1} p^k (1-p)^{(x-k)} & ; x = k, k+1, k+2, \ldots\ldots \\ 0 & ; d.d \end{cases}$$

biçiminde tanımlanır. Bu olasılık fonksiyonu k ve p değerlerine göre değiştiğinden dağılımın parametreleri k ve p'dir.

Dağılım ortalaması, varyansı ve M.Ç.F'u aşağıda verildiği gibidir.

$$E(X) = \mu = \frac{k}{p},\ Var(X) = \sigma^2 = \frac{kq}{p^2},\quad M_X(t) = \left(\frac{pe^t}{1-qe^t}\right)^k$$

İspat:

Burada $X = X_1 + X_2 + \ldots + X_k$ olarak yazılabilir. Bu toplamda yer alan $X_1, X_2, \ldots, X_k$ bağımsız, (aynı başarı olasılığı ile) geometrik dağılıma sahip değişkenlerdir.

Geometrik dağılım için aşağıdakilerin doğru olduğu daha önce gösterilmişti.

$$E(X_i) = \frac{1}{p};\ Var(X_i) = \sigma_{X_i}{}^2 = \frac{(1+q)}{p^2} - \left(\frac{1}{p}\right)^2 = \frac{q}{p^2}$$

$$M_{X_i}(t) = \frac{pe^t}{1-qe^t};\ i = 1,2,3,\ldots,k$$

$$\Rightarrow E(X) = \sum_{i=1}^{k} E(X_i) = \frac{k}{p};\ Var(X) = \sum_{i=1}^{k} Var(X_i) = \frac{kq}{p^2} \text{ ve}$$

$$M_X(t) = \prod_{i=1}^{k} M_{X_i}(t) = \left(\frac{pe^t}{1-qe^t}\right)^k \text{ olur.}$$

Örnek 5.6: Belirli bir bulaşıcı hastalığa maruz kalan bir çocuğun bu hastalığa yakalanma olasılığı 0.40 ise, bu hastalığa maruz kalan onuncu çocuğun bu hastalığa yakalanan üçüncü çocuk olma olasılığı nedir?

p= 0,40 =(bulaşıcı hastalığa yakalanma olasılığı)

x=10 =(hastalığa maruz kalan 10.uncu çocuk)

k=3 =(hastalığa yakalanan 3.cü çocuk)

$$P(x;k=3,p=0.40)=\binom{x-1}{2}(0.6)^{x-3}(0.4)^{3};\ x=3,4,5,...$$

$$=P(X=10;k=3,p=0.4)=\binom{9}{2}(0.4)^{3}(0.6)^{7}=(36)(0.4)^{3}(0.6)^{7}\cong 0.0645$$ olur.

Örnek 5.7: Bir makinenin ürettiği malların sağlam olması olasılığı 0.95'tir. Bu mallardan rastgele 12 tanesi alınıp denetleniyor. Denetlenen 12' inci malın 10' uncu sağlam mal olması olasılığı nedir?

$$P(x;k=10,p=0.95)=\binom{x-1}{9}(0.95)^{10}(0.05)^{x-10};$$ x= 10, 11, 12, ...

$$P(x=12;k=10,p=0.95)=\binom{11}{9}(0.95)^{10}(0.05)^{12-10}$$

$$=55(0.95)^{10}(0.05)^{2}\cong 0.08233$$

5.2.6. Hipergeometrik Dağılım

Binom dağılımında, her örnek çekimine ilişkin başarı olasılığı sabit ve örnek çekimlerinin sonuçları da birbirinden bağımsızdır. Binom dağılımını sonlu sayıda elemanı olan bir kümeden **iadeli** çekim gibi algılamak mümkündür.

Bazı durumlarda sonlu elemanı olan bir kümeden **iadesiz** örnek çekimi gerekebilir. Böyle hallerde, deney sonuçları (denemeler) birbirinden bağımsız olma özelliğini kaybeder.

Örnek çekiminin yapılacağı N elemanlı ana kitle içinde N_1 tanesi ilgilenilen sonucu (başarı), $N_2 = N - N_1$ tanesi ilgilenilmeyen sonucu (başarısız) verecek özellikteki elemanlara sahip alt kümeler olsun. N elemanlı bu ana kitleden n tanesi rastsal ve iadesiz olarak çekilip, çekilenler içinde kaç tanesinin 'Başarı' grubundan geleceği (X tanesinin) ile ilgileniyorsak, X'in dağılımı Hipergeometrik dağılım olarak adlandıracağımız bir dağılım modeline uyar.

Hipergeometrik dağılım aşağıdaki koşulların sağlanması halinde kullanılır:

- Her deneyin(çekimin) olası iki sonucu varsa,
- Deneyin tekrarlanma sayısı n sabitse
- Deneylerde birinin sonucu, diğerlerinin sonuçlarını etkiliyorsa (bağımlı deneyler ise)

Burada başarı grubundaki N_1 taneden x tanesinin seçimi $\binom{N_1}{x}$ değişik yolla, başarısızlar grubundaki N_2 taneden (n-x) tanesinin seçimi de $\binom{N_2}{n-x}$ değişik yolla yapılabilir. N elemanlı sonlu bir kümeden n tanesinin seçimi de $\binom{N}{n}$ değişik yolla yapılabilir. Buna göre yerine konulmaksızın seçilen n tane içindeki başarı sayısının x olması olasılığı: n, N_1 ve N_2 parametreleri ile Hipergeometrik dağılımın olasılık fonksiyonundan elde edilir.

$$P_X(x) = \frac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}} \quad x = 0,1,2,\ldots n, x \le N_1 \; ve \, (n-x) \le N_2$$

Hipergeometrik dağılımın ortalaması: $E(X) = \frac{nN_1}{N} = np$

Varyansı: $\mathrm{Var(X)} = n\frac{N_1}{N}\frac{N_2}{N}\frac{N-n}{N-1} = npq\frac{N-n}{N-1}$

olarak elde edilir. Yukarıdaki varyans formülünde $p = \frac{N_1}{N}$ ve $q = \frac{N_2}{N}$ $olarak$ algılanmalıdır.

Hipergeometrik dağılımının ortalaması (beklenen değeri) ile Binom dağılımın ortalaması aynıdır. Buna karşılık varyansları farklıdır. Hipergeometrik dağılım varyansındaki $\frac{N-n}{N-1}$ terimine düzeltme faktörü denir. $N \to \infty$ giderken düzeltme faktörü 1 olacağından Hipergeometrik dağılımın varyansı Binom dağılımına varyansına eşit olur.

Diğer yandan kitledeki N birimin N_1 ve N_2 gibi iki gruba değil de k tane gruba ayrıldığı düşünülürse bu durumda genelleştirilmiş Hipergeometrik dağılımdan yararlanılır.

$N = N_1 + ...N_k$ ve $n = x_1 + ... + x_k$ olmak üzere genelleştirilmiş Hipergeometrik dağılımın olasılık fonksiyonu:

$$P_X(x_1,.....x_k) = \frac{\binom{N_1}{x_1}\binom{N_2}{x_2}.....\binom{N_k}{x_k}}{\binom{N}{n}}$$ şeklinde tanımlanır.

Örnek 5.8: İçerisinde 3 kırmızı, 4 siyah ve 3 yeşil top bulunan bir kavanozdan yerine

Konulmaksızın 3 top seçiliyor.

a) İki tanesinin yeşil top olması olasılığını,

b) Birini yeşil, birini siyah ve birinin kırmızı top olması olasılığını bulunuz.

Cevaplar:

a) X rastlantı değişkeni seçilecek yeşil top sayısını göstersin. Kavanozda N=10 tane top var. Bu topları N_1 =3 yeşil olanlar ve $N-N_1=N_2=7$ yeşil olmayan diye iki gruba ayırırsak Hipergeometrik dağılımın olasılığını bulabiliriz.

$$P_X(x=2)=\frac{\binom{N_1}{x=2}\binom{N_2}{n-2}}{\binom{N}{n=3}}=\frac{\binom{3}{2}\binom{7}{1}}{\binom{10}{3}}=\frac{3(7)}{120}=\frac{7}{40}=0.175 \text{ bulunur.}$$

b) N tane top üç gruba ayrılacağından genelleştirilmiş Hipergeometrik dağılımdan yararlanılır.

$$P(X_1=1,X_2=1,X_3=1)=\frac{\binom{N_1}{x_1}\binom{N_2}{x_2}\binom{N_3}{x_3}}{\binom{N}{n}}=\frac{\binom{3}{1}\binom{4}{1}\binom{3}{1}}{\binom{10}{3}}=\frac{36}{120}=0.30 \text{ bulunur.}$$

Örnek 5.9: Bir işe başvuran 120 adaydan yalnız 80'i uygun niteliktedir. Ayrıntılı bir görüşme için adaylardan 5 tanesi rastgele seçildiğinde, bunlardan yalnızca 2'sinin uygun nitelikte olması olasılığı nedir?

N= 120 adaya başvurmuş N_1=80'i uygun nitelikte, N_2=40'ı uygun nitelikte değil

n=5 rastgele seçilen aday sayısı

X=2 rastgele seçilenlerden uygun olanların sayısı

$$P(X=2)=\frac{\binom{80}{2}\binom{40}{3}}{\binom{120}{5}}=\frac{(9880)(3160)}{190578024}\cong 0.164$$

5.2.7. Poisson Dağılımı

Araştırmacıların Binom dağılımı gibi sıkça kullandıkları dağılımlardan birisi de Poisson dağılımıdır. Binom dağılımı olasılıklarının bulunması n büyüdükçe zorlaşmaktadır. Bu nedenle Binom dağılımında $n \to \infty$, $p \to 0$ iken np sabit kaldığında Binom dağılımının limit durumu incelenirse, np=λ olmak üzere $p = \dfrac{\lambda}{n}$ için Binom dağılımının olasılık fonksiyonu

$$P(x;n,p) = \binom{n}{x}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x}$$ olarak ifade edilebilir.

$$\lim_{n\to\infty}\left[\binom{n}{x}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x}\right] = \frac{\lambda^x e^{-\lambda}}{x!}$$ olur.

Yukarıdaki limit sonucunun gerçeklemesi aşağıda verilmiştir.

$$\lim_{n\to\infty} P_x(x;n,p) = \lim_{n\to\infty}\left[\binom{n}{x}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x}\right]$$

$$= \lim_{n\to\infty}\left[\frac{n(n-1)\ldots(n-x+1)}{n^x}\frac{\lambda^x}{x!}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-x}\right]$$

$$= \frac{\lambda^x}{x!}\lim_{n\to\infty}\left[\frac{n(n-1)\ldots(n-x+1)}{\underbrace{n.n.\ldots.n}_{x\,\text{tane}}}\frac{\lambda^x}{x!}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-x}\right]$$

$$=\frac{\lambda^x}{x!}\lim_{n\to\infty}\left[\frac{1\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{x-1}{n}\right)}{\left(1-\frac{\lambda}{n}\right)^x}\left(1-\frac{\lambda}{n}\right)^n\right]$$

Parantez içindeki limitleri bulup yerine koyalım:

$$\lim_{n\to\infty}\left[1\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\ldots\left(1-\frac{x-1}{n}\right)\right]=1$$

$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^x=1$ ve $\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n=e^{-\lambda}$ olduğu için

$$\lim_{n\to\infty}P_x(x;n,p)=\lim_{n\to\infty}\left[\binom{n}{x}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x}\right]=\frac{\lambda^x}{x!}\frac{1}{1}e^{-\lambda}=\frac{\lambda^x.e^{-\lambda}}{x!}$$ elde edilir.

Buna göre n büyüdükçe ve p de ne kadar küçük ise Binom dağılımı yukarıda elde edilen dağılıma yaklaşır. Bu dağılıma Poisson dağılımı denir. λ Parametreli Poisson dağılımının olasılık fonksiyonu

$$P_X(x;\lambda)=P_X(x)=\frac{\lambda^x.e^{-\lambda}}{x!},x=0,1,2,3,\ldots.$$ Şeklinde tanımlanır.

Poisson dağılımının sadece Binom olasılıklarının hesabında kolaylık sağlayan bir dağılım olduğu düşüncesi doğru değildir. Poisson dağılımı, gerçekte sıkça karşılaşılan, olasılığa bağlı bazı süreçlerin modellenmesinde de çok önemli yere sahiptir.

Poisson olasılık fonksiyonunun kullanım alanı çok geniştir. Örneğin,

X= 'Bir radyoaktif kaynaktan bir zaman dilimi içinde yayılan alfa parçacıklarının Sayısı' olarak tanımlansa X 'n olasılık dağılımı, uygun bir μ değeri ile $P_X(x;\mu)=P_X(x)=\frac{\mu^x.e^{-\mu}}{x!},x=0,1,2,3,\ldots$ olasılık fonksiyonu ile modellenebilir.

Benzeri şekilde, X=' Bir üründeki kusur sayısı' olarak tanımlanırsa, X'in dağılımı, uygun bir m değeri ile, yine bir Poisson olasılık fonksiyonu ile modellenebilir.

X= ‘ bir zaman aralığında Meydana gelen trafik kazaları sayısı’ olarak tanımlanırsa, yine X’in dağılımı bir Poisson olasılık fonksiyonu ile modellenebilir.

Poisson dağılımının ortalama ve varyansı birbirine eşit olup

$E(X) = Var(X) = \lambda$ ’dır.

MÇF’nu ise $M_X(t) = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda} \sum \frac{(\lambda e^t)^x}{x!} = e^{-\lambda} e^{\lambda e^t} = e^{-\lambda(1-e^t)}$ olur.

İspat:

$$\mu = E(X) = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{(x-1)!},$$

$(x-1) = t$ olarak alınırsa $\Rightarrow x = 1$ *iken* $t = 0$ olur ve $x = t+1$ olur.

Bu değişken değişimi sonrası:

$\mu = E(X) = \sum_{t=0}^{\infty} \frac{e^{-\lambda}\lambda^{t+1}}{t!} = \lambda \sum_{t=0}^{\infty} \frac{e^{-\lambda}\lambda^t}{t!} = \lambda$ ve benzeri şekilde $E(X^2)$ hesaplanırsa:

$$E(X^2) = \sum_{x=0}^{\infty} x^2 \frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=1}^{\infty} x^2 \frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^x}{(x-1)!} = \sum_{t=0}^{\infty} (t+1) \frac{e^{-\lambda}\lambda^{(t+1)}}{t!} =$$

$$\lambda \sum_{t=0}^{\infty} t \frac{e^{-\lambda}\lambda^t}{t!} + \lambda \sum_{t=0}^{\infty} \frac{e^{-\lambda}\lambda^t}{t!} = \lambda E(X) + \lambda(1) = \lambda^2 + \lambda$$

Buradan $Var(X) = \sigma^2 = E(X^2) - \mu^2 = \lambda^2 + \lambda - (\lambda)^2 = \lambda$ bulunur.

Moment Çıkaran Fonksiyon:

$$M_X(t) = E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} = e^{-\lambda} e^{\lambda e^t} = e^{-\lambda + \lambda e^t}$$

***Örnek 5.11*:** Bir fabrikada depolanan ürünlerin 0,01'nin bozuk olduğu bilinmektedir. Bu depodan rastgele seçilen 50 birimden en az 4 tanesinin bozuk olma olasılıklarını hem Binom hem de Poisson dağılımından hesaplayalım.

X değişkeni n=50 lik bir örnekteki bozuk sayısı olsun.

Binom dağılımı ile n=50, p=0,01

$$P(X \geq 4) = 1 - P(X < 4) = 1 - \left[P(X=0) + P(X=1) + P(X=2) + P(X=3)\right]$$

$$= 1 - \left[\begin{array}{l} \binom{50}{0}(0.01)^0(0.99)^{50} + \binom{50}{1}(0.01)^1(0.99)^{49} + \\ \binom{50}{2}(0.01)^2(0.99)^{48} + \binom{50}{3}(0.01)^3(0.99)^{47} \end{array}\right]$$

$$= 1 - \left[\begin{array}{l} (0.99)^{50} + 50(0.01)(0.99)^{49} + \\ 1225(0.01)^2(0.99)^{48} + 19600(0.01)^3(0.99)^{47} \end{array}\right] \cong$$

$$1 - \left[0.605 + 0.3056 + 0.0756 + 0.0122\right] = 0.0016$$

Oldukça karmaşık olan bu işlemler yerine Poisson dağılımının olasılık fonksiyonundan yararlanılarak istenilen olasılık bulunabilir. $\lambda = np = 50(0.01) = 0.5$

$$= 1 - \left[\frac{e^{-0.5}(0.5)^0}{0!} + \frac{e^{-0.5}(0.5)^1}{1!} + \frac{e^{-0.5}(0.5)^2}{2!} + \frac{e^{-0.5}(0.5)^3}{3!}\right]$$

$$= 1 - \left[0.6065 + 0.30327 + 0.07582 + 0.01264\right] =$$

$$1 - 0.9982 = 0.0018$$

Poisson olasılık fonksiyonu aracılığı ile bulunan olasılık değeri gerçek olasılık değerine (Binom olasılık fonksiyonu ile hesaplanana) çok yakın çıkmıştır.

Örnek 5.12: Türkiye'de maden ocaklarında oluşan kazalar sonucunda her yıl ortalama olarak 1000 maden işçisinden bir tanesinin hayatını kaybettiği bilinmektedir. 2000 maden işçisinin çalıştığı bir maden ocağında bir yıl içinde,

a) sıfır işçinin

b) 3 işçinin

c) 2'den fazla işçinin

hayatını kaybetme olasılıklarını bulunuz.

X rastlantı değişkeni maden ocaklarında hayatını kaybeden işçi sayısı

n= 1000 iken E(X)=1=np $\Rightarrow$ p=0.001 olur

n= 2000 iken E(X) = np = $\lambda \Rightarrow \lambda$ =2000 (0.001) = 2 bulunur. (2000 işçiden bir yılda hayatını kaybeden işçi sayısı, ortalama olarak, 2 kişidir).

Cevap:

a) $P(X=0)=\dfrac{e^{-2}2^0}{0!}\cong 0.135$

b) $P(X=3)=\dfrac{e^{-2}.2^3}{3!}\cong 0.18045$

c) $P(X>2)=1-P(X\le 2)=1-\sum_{x=0}^{2}\dfrac{e^{-\lambda}\lambda^x}{x!}$

$$=1-\left[\frac{e^{-2}.2^0}{0!}+\frac{e^{-2}2^1}{1!}+\frac{e^{-2}.2^2}{2!}\right]=1-(0.135+0.2707+0.2707)$$

$\approx$1-0.6764=0.3236

Örnek 5.13: Ehliyet almak amacıyla sınava giren adayların %70'nin başarılı olduğu bilinmektedir. Ehliyet amacıyla sınava giren bir kişinin beşinci denemesinde başarılı olma olasılığı nedir?

X değişkeni =(ehliyet almak amacıyla girilen sınav sayısı) olsun.

P(X=5) olasılığı geometrik dağılımdan bulunur.

p=0.7olması halinde $P(X=x)=(0.3)^{x-1}p,\ x=1,2,3,\ldots$ olur ve buradan

$P(X=5)=(0.3)^{5-1}(0.7)=(0.3)^{4}(0.7)=0.0567$ bulunur.

Örnek 5.14: Otobüs ve kamyon üretimi yapan bir firmada günde ortalama 3 kamyon üretilmektedir. Bir günde üretilen kamyon sayısının 2'den az olması olasılığı nedir?

X değişkeni (bir günde üretilen kamyon sayısı) olsun.

λ = (bir günde üretilen ortalama kamyon sayısı) = 3 verildiğine göre

$P(X<2)$ Olasılığı Poisson dağılımından bulunur.

$$P(x<2)=\sum_{x=0}^{1}\frac{e^{-\lambda}\lambda^{x}}{x!}=\frac{e^{-3}.3^{0}}{0!}+\frac{e^{-3}.3^{1}}{1!}=0.0497+0.1494=0.1991 \text{ bulunur.}$$

Örnek 5.15: Bir yerleşim yerindeki nüfus içinde sol elini kullananların oranı %1 civarındadır. Rastgele seçilen 200 kişiden en fazla 196'sının sağ elini kullanması olasılığı nedir?

X değişkeni = (seçilen 200 kişi içinde sol elini kullananların sayısı) olsun.

P(200 kişiden en fazla 196'sını sağ elini kullanması)=

P(200 kişiden en az 4'ünün sol elini kullanması) demektir.

n= 200 p=0.01

$P(X \geq 4) = 1 - P(X < 4)$ olasılığı Binom yada Poisson dağılımı yardımı ile bulunabilir. (n=200 ve p=0.01 olduğundan)

Binom dağılımı ile

$$P(X \geq 4) = 1 - \sum_{x=0}^{3} \binom{200}{x} (0.01)^x (0.99)^{200-x}$$

$$= 1 - \left\{ \begin{array}{l} \binom{200}{0}(0.01)^0(0.99)^{200} + \binom{200}{1}(0.01)^1(0.99)^{199} \\ + \binom{200}{2}(0.01)^2(0.99)^{198} + \binom{200}{3}(0.01)^3(0.99)^{197} \end{array} \right\}$$

$$= 1 - 0.858 \cong 0.142$$

Poisson dağılımı ile

$\lambda = np = 200(0.01) = 2$

$$P(X \geq 4) = 1 - \sum_{x=0}^{3} \frac{e^{-2} 2^x}{x!} - 1 - \left\{ \frac{e^{-2} 2^0}{0!} + \frac{e^{-2} 2^1}{1!} + \frac{e^{-2} 2^2}{2!} + \frac{e^{-2} 2^3}{3!} \right\}$$

=1-0.857 $\cong$ 0.143 bulunur.

Not: n büyük ve p küçük olduğunda Binom dağılımı, Poisson dağılımına yakınsar.

Örnek 5.16: Bir makinenin ürettiği malların sağlam olması olasılığı 0.95'tir. Rastgele seçilen 12. ci Ürünün 10.uncu iyi ürün olması olasılığı nedir?

p=0.95, k=10, x=12 olan negatif Binom dağılımından yararlanılır.

$$P(x;k,p)=\binom{x-1}{k-1}p^{k}(1-p)^{x-k}\quad x=k,k+1,k+2,.....$$

$$P(X=12;k=10;p=0.95)=\binom{11}{9}(0.95)^{10}(0.05)^{2}\cong 0.082$$