Kümülatif Dağılım Fonksiyonları

- Herhangi bir rastgele değişken için kümülatif dağılım fonksiyonu/cumulative distribution function (KDF/CDF) şu şekilde tanımlanır.

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(x')\, dx'$$

- Sürekli olan içinse

$$P_X(x) = P(X \leq x) = \sum_{x'=-\infty}^{x} p_X(x')$$

- KDF bir olasılık olduğundan değeri hep sıfır ile bir arasındadır.
- Ayrıca tanımdan dolayı -sonsuz'da sıfır, +sonsuz'da ise bir değerini alır.
- KDF ayrıca hep monoton artan veya sabit bir fonksiyoundur, yani artan x değeri ile hiçbir şekilde azalması mümkün değildir.
- Verilmiş olan bir KDF için OYF türev işlemi ile bulunur. Yani:

$$f_X(x) = \frac{dF_X(x)}{dx}$$

Çıkarılmış Olasılık Yoğunluk Fonksiyonları

- Kumulatif dağılım fonksiyonlarının en önemli kullanım alanlarından biri ”çıkarılmış olasılık yoğunluk fonksiyonu (derived PDF)” yönteminde ortaya çıkar
- Çıkarılmış OYF Metodu, OYF’si verilmilş bir RD’nin bir fonksiyonu olan başka bir RD verildiğinde, bu yeni RD’nin OYF’sini hesaplamanın sistematik bir yöntemidir.
- Yani çıkarılmış OYf/derived PDF, bir x RDi ve OYF’si $f_X(x)$ verilmiş olduğunda, bize verilen $y = g(x)$ için y’nin OYFsi $f_Y(y)$’yi hesaplamanın yöntemidir.

Çıkarılmış Olasılık Yoğunluk Fonksiyonları Aşamaları

- Çıkarılmış OYF Yönteminini üç aşaması vardır ve şu şekilde açıklanabilir.
- Problem: x ve $f_X(x)$ verilmiş olsun; $y_g(x)$ (örneğin $y = \pi x^2$ veya $y = \sin(x)$ vb) için $f_Y(y)$ nedir?
- Aşama 1: $f_X(x)$ kullanılarak $F_X(x)$ hesaplanır: $F_X(x) = \int_{-\infty}^{x} f_X(x')\,dx'$
- Aşama 2: $y = g(x)$ bağıntısı kullanılarak $F_Y(y)$ hesaplanır: $F_Y(y) = P(Y \leq y) = P(g(x) \leq y)$. Burada $g(x)$'in formuna göre x yalnız bırakılarak $F_Y(y)$, $F_X(x)$ kullanılarak yazılır.
- Aşama 3: $F_Y(y)$'nin türevi alınarak ($F_X(x)$ de kullanılarak) $f_Y(y)$ bulunur.
- Bu konu ile ilgili birçok örnek derste çözülmüştür.

Kovaryans (Covariance) ve Korelasyon Katsayısı (Correlation Coefficient)

- İki X ve Y gibi RD'nin kovaryansı şu şekilde hesaplanır:

$$[\mathrm{Cov}](\mathrm{x},\mathrm{y}) = \mathrm{E}[\mathrm{X} - \mu_{\mathrm{x}}]\mathrm{E}[\mathrm{Y} - \mu_{\mathrm{y}}]$$

- Kovaryans iki RD'nin birbirini etkileme/birbiri ile ilişkisi miktarını ifade eder.
- Korelasyon Katsayısı (Correlation Coefficient) ise şu şekilde tanımlanır:

$$\rho(X,Y) = \frac{[\mathrm{Cov}](\mathrm{x},\mathrm{y})}{\sigma_x \sigma_y}$$

- Korelasyon katsayısının kovaryanstan farkı, hep 0 ile 1 arasında bir sayı verecek şekilde normalize edilmiş olmasıdır.
- Böylece verilmiş olan bir korelasyon katsayısına bakarak, 0 veya 1'e yakınlığını baz alabilir ve ili RD arasındaki ilinti ya da ilintisizliğin ne kadar kuvvetli olduğunu anlayabiliriz.
- Kovaryansı sıfır olan iki RD'ye "ilintisiz" denir. Bağımsız olan iiki RD kesinlikle ilintisizdir ama tersi doğru değildir.

Yenilemeli Beklenen Değer ve Varyans

- Birbirine bağlı iki rastgele değişkenin olduğu durumlarda, beklenen değer ve varyans şu şekilde iki aşamada hesaplanabilir.

$$\mathrm{E}[X] = \mathrm{E}_Y[\mathrm{E}_X[X/Y]$$

- Burada ilk beklenen değer verilmiş olan bir Y için Xe göre, ikinci beklenen değer ise Yye göredir.
- Benzer şekilde varyans da şöyle hesaplanabilir.

$$\mathrm{Var}(X) = \mathrm{E}[\mathrm{Var}(X/Y)] + \mathrm{Var}(\mathrm{E}[X/Y])$$

- Bu iki formülde amaç başka bir RD olan Yye bağlı olduğu için Xin beklenen değeri ve varyansını direkt olarak hesaplayamadığımız durumlarda problemi ikiye bölmektir. İlk kısımde Y sabit kabul edilir, ikinci kısımda ise Y üzerinden hesaplamalar yapılır.

Örnek

- l uzunluğunda bir çubuğu düzgün dağılımlı olarak rastgele biryerinden kırıyoruz. Sol elimizde kalan parçanın uzunluğuna Y diyelim. Sonra bu kalan çubuğu tekrar düzgün dağılımlı olarak kırıyoruz. İkinci kırmadan sonra en son elimizde kalan parçanın uzunluğu ise X olsun. Xin beklenen değerş ve varyansı nedir?
- Burada beklenen değer ve varyans için basit tanım formüllerini kullanarak hesap yapamıyoruz, çünkü X başka bir RD olan Yye bağlı. O zaman bir önceki slide'da yer alan yaklaşımı ve formülleri kullanmamız gerekiyor

$$\mathrm{E}[X] = \mathrm{E}_Y[\mathrm{E}_X[X/Y] = \mathrm{E}_Y[Y/2] = l/4$$

- Burada $\mathrm{E}_X[X/Y] = Y/2$ oldu çünkü sabir bir verilmiş Y değeri için ortalama $Y/2$dir (düzgün dağılım olduğu için).
- Daha sonra ikinci beklenen değeri Yye göre alıyoruz: $\mathrm{E}_Y[Y/2] = 1/2 \times E[Y] = 1/2 \times l/2 = l/4$
- Varyans ise şöyle hesaplanıyor:

$$\mathrm{Var}(X) = \mathrm{E}[\mathrm{Var}(X/Y)] + \mathrm{Var}(\mathrm{E}[X/Y]) = \mathrm{E}[Y^2/12] + \mathrm{Var}(Y/2) = l^2/36 + l^2/48 = 7l^2/144$$