OLASILIKVE İSTATİSTİK

BAZI SÜREKLİ DAĞILIMLAR

Düzgün Dağılım

X rasgele değişkeninin olasılık yoğunluk fonksiyonu

$$f(x) = \begin{cases} \dfrac{1}{b-a} & , \quad a \leq x \leq b \\ 0 & , \quad d.d. \end{cases}$$

ise X rasgele değişkeni $[a, b]$ kapalı aralığında *sürekli düzgün dağılıma* sahiptir denir.

X rasgele değişkeni sürekli düzgün dağılıma sahipse,

$$E(X) = \frac{a+b}{2}$$

$$Var(X) = \frac{(b-a)^2}{12}$$

Örnek.

E1 numaralı otobüs, Denizevleri durağına saat 10:00 ile 10:30 arasında düzgün dağılıma uygun herhangi bir zamanda gelmektedir. Saat 10:00 da durakta olduğumuzu varsayalım.

a) E1' in gelmesini 10 dakikadan fazla bekleme olasılığımız nedir?
b) Eğer saat 10:15 olmuş ve hala otobüs gelmemişse, en az 10 dakika daha bekleme olasılığımız nedir?

Çözüm. Düzgün dağılım söz konusu.

$$X: E1 \text{ otobüsünü bekleme süresi}$$

$$f(x) = \begin{cases} \dfrac{1}{b-a} & , \quad a \leq x \leq b \\ 0 & , \quad d.d. \end{cases}$$

$$f(x) = \begin{cases} \dfrac{1}{30} & , \quad 10{:}00 \leq x \leq 10{:}30 \\ 0 & , \quad d.d. \end{cases}$$

Kafamız karışmaması için, 10:00 – 10:30 arasını 0-30 alalım. Yani 30 dakikalık zaman.

$$f(x) = \begin{cases} \dfrac{1}{30} & , \quad 0 \le x \le 30 \\ 0 & , \qquad d.d. \end{cases}$$

a) $P(X > 10) = \int_{10}^{30} \frac{1}{30} dx = \frac{20}{30} = \frac{2}{3} = 0.666$

b) $P(15 \le X \le 25) = \int_{15}^{25} \frac{1}{30} dx = \frac{10}{30} = \frac{1}{3} = 0.333$

Normal Dağılım

İstatistikte uygulanan birçok analiz ve tahmin yöntemleri X değişkeninin Normal dağılış göstermesi halinde geçerlidir. Dolayısıyla bu dağılış istatistikte önemli bir yere sahiptir.

Eldeki değişken sürekli ve tek modlu ise Normal dağılış göstermese de bazı dönüşümlerle dağılışı Normal dağılışa yaklaştırmak mümkün olabilir. Bunun için $\sqrt{X}$, $log(X)$, $1/X$ gibi dönüşümler kullanılabilir.

1809 tarihinde Alman Matematikçi Gauss tarafından bulunmuştur ve Gauss dağılımı olarak da bilinmektedir.

Normal dağılımın grafiğine normal eğri denir.

İstatistiksel analiz tekniklerinin büyük kısmı, anakütle dağılımının normal dağılıma uyduğu varsayımına dayalıdır.

X sürekli rasgele değişkeni için olasılık yoğunluk fonksiyonu

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2} \quad ; \quad -\infty < \mu < \infty , -\infty < X < \infty , \sigma^2 > 0$$

ise X rasgele değişkeni Normal dağılıma sahiptir denir. Burada

μ : Normal dağılımın ortalaması

σ^2 : Normal dağılımın varyansı

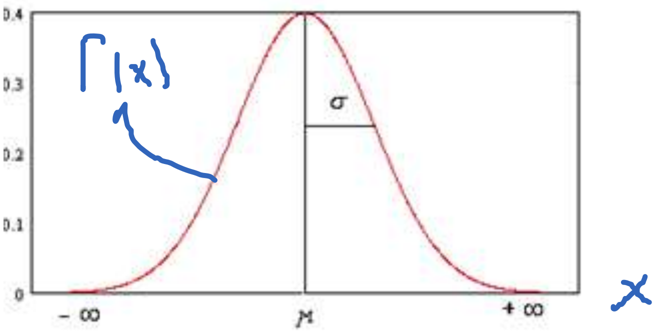
e : 2,371…

π : 3,14…..

dir.

X rasgele değişkeni, μ ortalamalı ve σ^2 varyanslı normal dağılıma sahipse $X \sim N(\mu, \sigma^2)$ ile gösterilir.

Normal dağılımın grafiği aşağıdaki gibidir.



$f(x)$ eğrisi altında kalan alan 1'e eşittir. Yani,

$$\int_{-\infty}^{\infty} f(x)\, dx = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2} dx = 1$$

dir.

Normal dağılımın özellikleri:

a) Normal dağılım $X = \mu$ doğrusuna göre simetriktir.

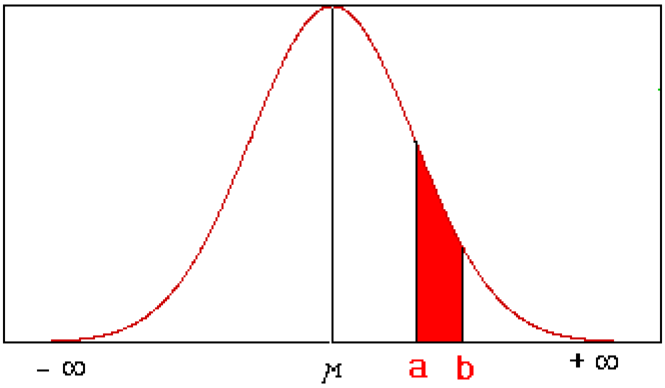
b) Ortalama, mod ve medyan birbirine eşittir. Dolayısıyla tümü aynı nokta üzerindedir.
c) Verilerin

yaklaşık %68.26 sı ortalama ± 1 standart sapma uzaklıkta,

yaklaşık %95.44 ü ortalama ± 2 standart sapma uzaklıkta,

yaklaşık %99.7 si ortalama ± 3 standart sapma uzaklıktadır.

$X \sim N(\mu, \sigma^2)$ iken X' in a ile b arasında olma olasılığı:



$$P(a < X < b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2} dx$$

ile bulunur. Bu integrali hesaplamak zaman alıcı ve zor olduğu için X rasgele değişkeni standartlaştırılarak standart normal dağılıma geçilir.

Standartlaştırma,

$$z = \frac{X-\mu}{\sigma}$$

ile yapılır.

Standart Normal Dağılım

Ortalaması sıfır $(\mu = 0)$, varyansı 1 $(\sigma^2 = 1)$ ve olasılık yoğunluk fonksiyonu

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad , \quad -\infty < z < \infty$$

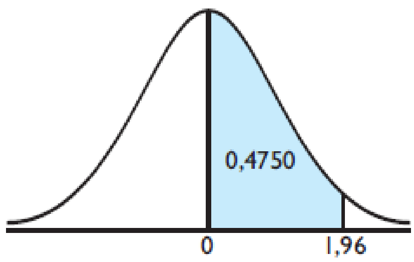
olan dağılıma standart normal dağılım denir.

$z = \frac{X-\mu}{\sigma}$ değişkeni standart normal dağılıma sahiptir. Yani $z \sim N(0,1)$ dir.

z değişkeni ile ilgili olasılıklar hesaplanırken, standart normal dağılım tablosu ($z\ tablosu$) kullanılır.

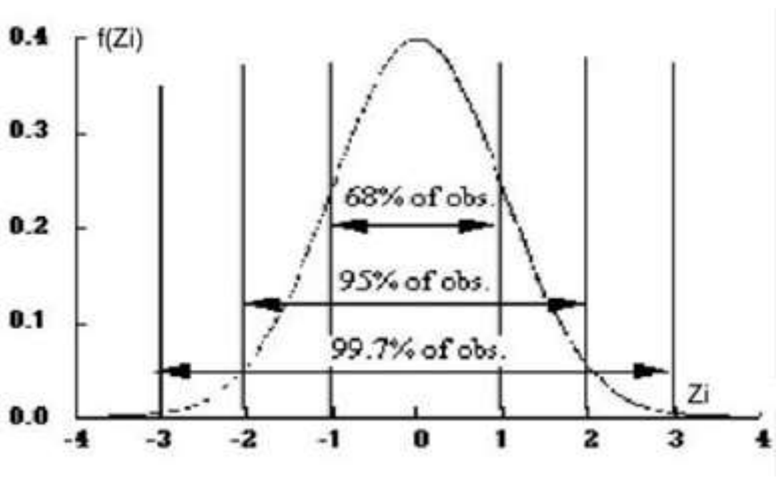
Ek-1: Standart Normal Dağılım Altında Kalan Alan (z) Tablosu

Eğer z=1,96 ise
P(0'dan z'ye)=0,4750



z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,00	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,10	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,20	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,30	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,40	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,50	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,60	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,70	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,80	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,90	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,00	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,10	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1,20	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

1,30	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,40	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,50	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,60	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,70	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,80	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,90	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,00	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,10	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,20	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,30	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,40	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,50	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,60	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,70	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,80	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,90	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,00	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990



Örnek: $X \sim N(3,4)$ ise X'in 3 ile 5 arasında bulunma olasılığı nedir?

Çözüm: $P(3 < X < 5) = \int_3^5 \frac{1}{3\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-3}{2}\right)^2} dx$

işleminin sunucu, istenen olasılığı verecektir.

Bunun yerine,

$z = \frac{X-\mu}{\sigma}$ dönüşümü yapıldığında ($z \sim N(0,1)$)

$$P(3 < X < 5) = P\left(\frac{3-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{5-\mu}{\sigma}\right) = P\left(\frac{3-3}{2} < z < \frac{5-3}{2}\right) = P(0 < z < 1)$$

olur.

$z\ tablosu$ kullanılarak,

$P(0 < z < 1) = 0.3413$

bulunur.