4.4 Cumulative Density Function (cdf)

The cumulative distribution function (CDF) calculates the cumulative probability for a given x-value. Use the CDF to determine the probability that a random observation that is taken from the population will be less than or equal to a certain value. You can also use this information to determine the probability that an observation will be greater than a certain value, or between two values.

4.5 Parameters

Parameters are descriptive measures of an entire population that may be used as the inputs for a probability distribution function (PDF) to generate distribution curves. Parameters are usually signified by Greek letters to distinguish them from sample statistics. For example, the population mean is represented by the Greek letter mu (μ) and the population standard deviation by the Greek letter sigma (σ). Parameters are fixed constants, that is, they do not vary like variables. However, their values are usually unknown because it is infeasible to measure an entire population.

4.6 Discrete Probability Distributions

4.6.1 Bernoulli Distribution

A random experiment of which the outcome can be classified into two categories is called a Bernoulli trial. In general, the results of a Bernoulli trial are called 'success' and 'failure'.

Examples:

- You take a pass-fail exam. You either pass or fail.
- You toss a coin. The outcome is ether heads or tails.
- A child is born. The gender is either male or female.

Consider a Bernoulli trial and let

$$X = \begin{cases} 0 & \text{if the Bernoulli trial results in a failure} \\ 1 & \text{if the Bernoulli trial results in a success} \end{cases}$$

Suppose that the probability of a 'success' in any Bernoulli trial is θ. Then X is said to have a Bernoulli distribution with probability mass function (pmf)

$$f_X(x) = \theta^x (1-\theta)^{1-x}\ ; x = 0,1$$

This is called as X~Bernoulli(θ).

Example: Namal and Laman went to the canteen and thought to buy a cake or a bread to eat. To decide the item, they tossed a fair coin. What is the probability of buying a cake?

4.6.2 Binomial Distribution

An experiment that conforms to the following list of requirements is called a 'binomial experiment'.

1. The experiment consists of a sequence of n trials, where n is fixed in advance of the experiment.
2. Each trial can result in one of the same two possible outcomes which we call 'success' (S) or 'failure' (F).
3. The trials are independent, so that the outcome of any particular trial does not influence the outcome of any other trial.
4. The probability of 'success' is the same for each trial; suppose this probability is θ.

Consider a binomial experiment with n trials and probability of θ success. Let,

$X =$ The number of 'successes' in n trials

Then, X is said to have a binomial distribution with probability mass function

$$f_X(x) = \binom{n}{x}\theta^x(1-\theta)^{1-x} \quad ; \quad x = 0,1,2,\dots,n$$

This is denoted as $X \sim Bin(n,\theta)$

Activity 4.2

Jeremy sells a magazine which is produced in order to raise money for homeless people. The probability of making a sale is, independently, 0.09 for each person he approaches. Given that he approaches 40 people, find the probability that he will make:

(a) 2 or fewer sales;
(b) exactly 4 sales;

4.6.3 Poisson Distribution

The Poisson distribution often provides a realistic probability model for the number of events in a given period of time or space. For example, probability concerning the number of traffic accidents per week in a given city, the number of radioactive particle emissions per unit of time, the number of telephone calls coming into a switchboard of a company per hour, the number of meteorites that collide with a test satellite during a single orbit, the number of bacteria per unit volume of some fluid, the number of defects per unit length of some material, the number of flaws per unit length of some wire, etc. can often be calculated using the Poisson distribution.

Let X be the number of events during a time interval. Suppose that the average number of events during the interested time interval is λ (>0). Then, the distribution of X can be modeled by a Poisson distribution with the p.m.f

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

This is denoted as $X \sim Poisson(\lambda)$

Activity 4.3

Births in a hospital occur randomly at an average rate of 1.8 births per hour.

(1) What is the probability of observing 4 births in a given hour at the hospital?

(2) What is the probability of observing less than 2 births in a given half an hour?

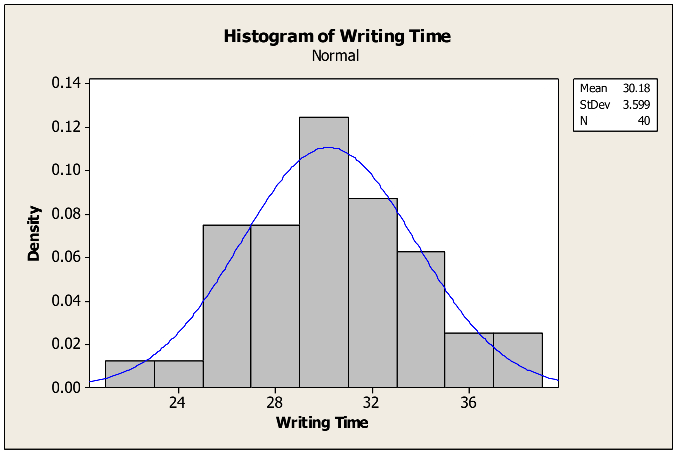
4.7 Continuous Probability Distributions

4.7.1 Normal Distribution

A consumer organization wishes to study about lifetime (hours of writing) of a brand of a ballpoint pen. The organization purchased 40 such pens and inserted each in a specially constructed machine that causes the pen to write continuously until the ink runs out. Table below gives the observed life times. Let X be the writing time of a randomly chosen pen. Suggest a suitable model for the pdf of X. Estimate $\Pr(30 \leq X \leq 35)$ using the model pdf.

31.7	26.2	35.0	34.8	33.1	30.8	34.1	34.4	31.4	29.5
29.0	29.7	32.5	25.3	30.6	37.7	28.7	23.1	25.7	26.7
31.0	28.7	30.7	36.4	27.3	32.1	29.2	30.1	26.3	25.9
27.5	28.6	37.2	31.3	22.7	27.4	31.7	30.0	33.6	29.6

Table of Hours of writing



Histogram of writing time data and a possible model for the pdf

It seems that a bell shaped curve is suitable as the model for pdf of X.

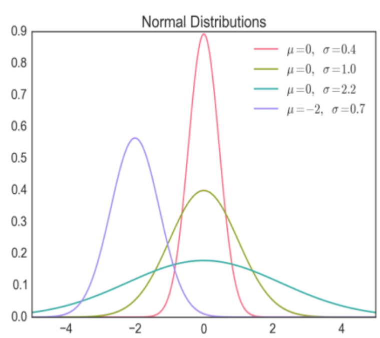
One commonly used bell shaped curve is called the normal distribution. It is probably the most oftenly used probability density function in the statistics world.

Random variable X is said to have a normal distribution with location parameter μ and scale parameter σ if its probability function is given by,

$$f_X(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\};\ -\infty < x < \infty$$

This is denoted by $X \sim N(\mu, \sigma^2)$

The normal density function is symmetric around the location parameter μ. The dispersion of the distribution depends on the scale parameter σ



4.7.2 Standard Normal Distribution

Normal distribution with $\mu = 0$ and $\sigma = 1$ is called the standard normal distribution. A random variable with a standard normal distribution is usually denoted by Z. The probability density function of standard normal distribution is denoted by φ.

If $Z \sim N(0,1)$, then

$$\Phi_Z(z) = \frac{1}{\sqrt{2\pi}} exp\left\{-\frac{1}{2}z^2\right\}; \quad -\infty < z < \infty$$

Probabilities related to Z can be found by using standard normal probability table.

Example

Let Z be a standard normal variable. Calculate probabilities given below.

No	Calculate the probability	Answer
1	Pr(Z>0.95)	
2	Pr(Z<2.02)	
3	Pr(-2.41<Z<-0.94)	
4	Pr(Z>-1.48)	
5	Pr(Z<-1.76)	
6	Pr(Z<1.7)	
7	Pr(Z<-0.33)	
8	Pr(0.06<Z<1.62)	
9	Pr(-2.49<Z<-1.42)	
10	Pr(Z>2.67)	
11	Pr(Z<1.75)	0.9599
12	Pr(Z>-1.3)	0.9032
13	Pr(Z>1.38)	0.0838
14	Pr(Z<-1.01)	0.1562
15	Pr(-2.96<Z<1.05)	0.8516
16	Pr(Z<1.88)	0.9699
17	Pr(Z>1.02)	0.1539
18	Pr(-2.07<Z<-0.75)	0.2074

19	Pr(Z>-2.27)	0.9884
20	Pr(Z>-0.77)	

Standard Normal Distribution Table (Right-Tail Probabilities)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002