MAT2377

Ali Karimnezhad

Version September 30, 2015

Comments

- These slides cover material from Chapter 3.
- In class, I may use a blackboard. I recommend reading these slides before you come to the class.
- I am planning to spend 4 lectures on this chapter.
- I am not re-writing the textbook. The reference book contains many interesting and practical examples.
- There may be some typos. The final version of the slides will be posted $after$ the chapter is finished.

Bernoulli Distribution

A Bernoulli trial is a random experiment with two possible outcomes, success and failure.
One obvious application deals with the testing of items as they come off an assembly line. Each trial may indicate a defective or a nondefective item.

- Denote p as the probability of a success.
- Let X be a sucess indicator in a Bernoulli trial.
- Then, $X \sim Ber(p)$ with the probability function

$$\begin{aligned} P(X=x) &= \begin{cases} 1-p & \text{for } x=0 \\ p & \text{for } x=1 \end{cases} \\ &= p^{x}(1-p)^{1-x}, \quad x=0,1. \end{aligned}$$

Binomial Distribution

A binomial experiment consists of n repeated independent Bernoulli trials, each with the same probability of succes.

- Denote p as the probability of a success.
- Let X be number of sucesses in a Binomial trial.
- Then, $X \sim B(n, p)$ with the probability function

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

- If $n = 1$, $X \sim B(1, p)$ or equivalently $X \sim Ber(p)$.

Examples:

- The probability that a patient recovers from a rare blood disease is 0.4. If 5 people are known to have contracted this disease, what is the probability that exactly 2 survive?

- A large chain retailer purchases a certain kind of electronic device from a manufacturer. The manufacturer indicates that the defective rate of the device is 3%. The inspector randomly picks 20 items from a shipment. What is the probability that there will be at most one defective item among these 20?

Expectation and Variance for Binomial Distribution

- From math basis, we know

$$(p+q)^n = \binom{n}{0} p^0 q^n + \binom{n}{1} p^1 q^{n-1} + \ldots + \binom{n}{n} p^n q^0.$$

- Expecation;

$$\mathrm{E}(X) = \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} = \cdots = np.$$

- Variance

$$\mathrm{Var}(X) = \mathrm{E}\left[(X-np)^2\right] = \sum_{x=0}^{n} (x-np)^2 \binom{n}{x} p^x (1-p)^{n-x} = np(1-p)\,.$$

Example:

Suppose that each sample of water has 10% of being polluted. If 12 samples are selected independently, then it is reasonable to model the number X polluted sample as in the sample as $B(12, 0.1)$. Find

(a) $\mathrm{E}(X)$ and $\mathrm{Var}(X)$.

(b) $P(X = 3)$.

(c) $P(X \leq 3)$ (tables can be used).

Solution:

(a) If $X \sim B(n,p)$ then $\mathrm{E}(X) = np$ and $\mathrm{Var}(X) = np(1-p)$ so

$$\mathrm{E}(X) = 12 \times 0.1 = 1.2 \quad ; \quad \mathrm{Var}(X) = 12 \times 0.1 \times 0.9 = 1.08\,.$$

(b) $P(X = 3) = \binom{12}{3}(0.1)^3(0.9)^9 \approx 0.0852.$

(c) $P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = \cdots.$
However, $P(X \leq 3)$ for $X \sim B(12, 0.1)$ is tabulated in Appendix A.1 of the Reference Book: $\approx 0.9744.$

- Also, since $P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) \approx 0.8891$ from tables we see that $P(X = 3) = P(X \leq 3) - P(X \leq 2) \approx 0.9744 - 0.8891 = 0.0853$ (note rounding error).

Table A.1 (continued) Binomial Probability Sums $\sum_{x=0}^{r} b(x; n, p)$

		p									
n	***r***	**0.10**	**0.20**	**0.25**	**0.30**	**0.40**	**0.50**	**0.60**	**0.70**	**0.80**	**0.90**
12	**0**	0.2824	0.0687	0.0317	0.0138	0.0022	0.0002	0.0000			
	1	0.6590	0.2749	0.1584	0.0850	0.0196	0.0032	0.0003	0.0000		
	2	0.8891	0.5583	0.3907	0.2528	0.0834	0.0193	0.0028	0.0002	0.0000	
	3	0.9744	0.7946	0.6488	0.4925	0.2253	0.0730	0.0153	0.0017	0.0001	
	4	0.9957	0.9274	0.8424	0.7237	0.4382	0.1938	0.0573	0.0095	0.0006	0.0000
	5	0.9995	0.9806	0.9456	0.8822	0.6652	0.3872	0.1582	0.0386	0.0039	0.0001
	6	0.9999	0.9961	0.9857	0.9614	0.8418	0.6128	0.3348	0.1178	0.0194	0.0005
	7	1.0000	0.9994	0.9972	0.9905	0.9427	0.8062	0.5618	0.2763	0.0726	0.0043
	8		0.9999	0.9996	0.9983	0.9847	0.9270	0.7747	0.5075	0.2054	0.0256
	9		1.0000	1.0000	0.9998	0.9972	0.9807	0.9166	0.7472	0.4417	0.1109
	10				1.0000	0.9997	0.9968	0.9804	0.9150	0.7251	0.3410
	11					1.0000	0.9998	0.9978	0.9862	0.9313	0.7176
	12						1.0000	1.0000	1.0000	1.0000	1.0000
13	**0**	0.2542	0.0550	0.0238	0.0097	0.0013	0.0001	0.0000			
	1	0.6213	0.2336	0.1267	0.0637	0.0126	0.0017	0.0001	0.0000		
	2	0.8661	0.5017	0.3326	0.2025	0.0579	0.0112	0.0013	0.0001		
	3	0.9658	0.7473	0.5843	0.4206	0.1686	0.0461	0.0078	0.0007	0.0000	
	4	0.9935	0.9009	0.7940	0.6543	0.3530	0.1334	0.0321	0.0040	0.0002	
	5	0.9991	0.9700	0.9198	0.8346	0.5744	0.2905	0.0977	0.0182	0.0012	0.0000
	6	0.9999	0.9930	0.9757	0.9376	0.7712	0.5000	0.2288	0.0624	0.0070	0.0001
	7	1.0000	0.9988	0.9944	0.9818	0.9023	0.7095	0.4256	0.1654	0.0300	0.0009
	8		0.9998	0.9990	0.9960	0.9679	0.8666	0.6470	0.3457	0.0991	0.0065
	9		1.0000	0.9999	0.9993	0.9922	0.9539	0.8314	0.5794	0.2527	0.0342
	10			1.0000	0.9999	0.9987	0.9888	0.9421	0.7975	0.4983	0.1339
	11				1.0000	0.9999	0.9983	0.9874	0.9363	0.7664	0.3787
	12					1.0000	0.9999	0.9987	0.9903	0.9450	0.7458
	13						1.0000	1.0000	1.0000	1.0000	1.0000

Example:

According to a chemical engineering progress study, approximately 30% of all pipework failures in chemical plants are caused by operator error.

(a) What is the probability that out of the next 13 pipework failures at least 10 are due to operator error?

(b) What is the probability that no more than 4 out of 13 such failures are due to operator error?

Multinomial Distribution

The Binomial experiment becomes a Multinomial experiment if we let each trial have more than two possible outcomes.

If a given trial can result in the k outcomes $E_1, E_2, \ldots, E_k$ with probabilities $p_1, p_2, \ldots p_k$, then the probability distribution of the random variables $X_1, X_2, \ldots, X_k$, representing the number of occurrences for $E_1, E_2, \ldots, E_k$ in n independent trials, is

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k) = \binom{n}{x_1, x_2, \ldots, x_k} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k},$$

where $x_1 + x_2 + \ldots + x_k = n$ and $p_1 + p_2 + \ldots + p_k = 1$.

Example:

In a certain town, 40% of the eligible voters prefer candidate A, 10% prefer candidate B, and the remaining 50% have no preference. You randomly sample 10 eligible voters. What is the probability that 4 will prefer candidate A, 1 will prefer candidate B, and the remaining 5 will have no preference?

Hypergeometric Distribution

Example-Motivation:

- Suppose that an urn contains 8 red balls and 4 green balls. We draw 2 balls from the urn without replacement. If we assume that at each draw each ball in the urn is equally likely to be chosen, what is the probability that both balls are Red?
  Define $R_1 = \{\text{the first ball is Red}\}$ and $R_2 = \{\text{the second ball is Red}\}$:

$$P(R_1 \cap R_2) = \frac{\binom{8}{2}}{\binom{12}{2}} = \ldots = \frac{8 \times 7}{12 \times 11}.$$

  Let X be number of Red balls in the drawing:

$$P(X = 2) = \frac{\binom{8}{2}\binom{4}{0}}{\binom{12}{2}} = \ldots = \frac{8 \times 7}{12 \times 11}.$$

Hypergeometric Distribution

A Hypergeometric experiment possesses the following two properties:

1. A random sample of size n is selected without replacement from N items.

2. Of the N items, k interesting items have a same property in common, and $N-k$ have another same property in common.

If X is number of type-k items in a random sample of size n, then $X \sim HG(N,n,k)$ with the probability function

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

Example:

Lots of 40 components each are deemed unacceptable if they contain 3 or more defectives. The procedure for sampling a lot is to select 5 components at random and to reject the lot if a defective is found.
What is the probability that exactly 1 defective is found in the sample if there are 3 defectives in the entire lot?

Define X as number of defective items. Then, using the Hypergeometric distribution with $n = 5$, $N = 40$ and $k = 3$, $X \sim HG(40, 5, 3)$. Thus,

$$P(X = 1) = \frac{\binom{3}{1}\binom{40-3}{5-1}}{\binom{40}{5}} = 0.3011.$$

Expectation and Variance for Hypergeometric Distribution

If $X \sim HG(N, n, k)$,

- Expecation

$$\mathrm{E}(X) = \sum_{x=0}^{n} x \frac{\binom{k}{x}\binom{n-k}{n-x}}{\binom{N}{n}} = \frac{nk}{N}.$$

- Variance

$$\mathrm{Var}(X) = \mathrm{E}\left[(X - \frac{nk}{N})^2\right] = \frac{N-n}{N-1}\frac{nk}{N}(1 - \frac{k}{N}).$$

For a proof, see Appendix A.12 of the reference book.

Geometric Distribution

Example-Motivation:

- Flip a coin repeatedly until you get a Head. Let the probability that the coin lands Heads up be p and the probability that the coin lands Tails up be $q = 1 - p$. Suppose at each flip a Head or a Tail is equally likely to be observed.
  What is probability of observing a Head (success) in the 4th flip?

  Define X as number of required trails to observe a Head (success). Thus,

$$P(X = 4) = P(\{TTTH\}) = p\,q^3.$$

Geometric Distribution

A Geometric experiment possesses the following two properties:

1. A Bernoulli trail with probability p of success and probability of failure $q = 1 - p$ at each step is repeated.

2. The experiment is stopped when only one success occurs.

If X is the number of steps required until the 1st success occurs, then $X \sim Ge(p)$ with the probability function

$$g(x;p) = P(X = x) = p\,q^{x-1}, \quad x = 1, 2, \ldots.$$

Example:

For a certain manufacturing process, it is known that, on the average, 1 in every 100 items is defective.

What is the probability that the 5th item inspected is the first defective item found?

Define X as number of required trails until the 1st defective item is found. Then, using the Geometric distribution with $p = \frac{1}{100}$, $X \sim Ge(0.01)$. Thus,

$$P(X = 5) = (0.01)\,(0.99)^4 = 0.0096.$$

Expectation and Variance for Geometric Distribution

If $X \sim Ge(p)$,

- Expectation

$$\mathrm{E}(X) = \sum_{x=1}^{\infty} x\, pq^{x-1} = \frac{1}{p}.$$

- Variance

$$\mathrm{Var}(X) = \mathrm{E}\left[(X - \frac{1}{p})^2\right] = \frac{q}{p^2}.$$

Negative Binomial Distribution

Example-Motivation:

- In a game, we flip a coin repeatedly until we get two Heads (the game is stopped when the second Head is observed). Let the probability that the coin lands Heads up be p and the probability that the coin lands Tails up be $q = 1 - p$. Suppose at each flip a Head or a Tail is equally likely to be observed.

  What is probability that the game is stopped in the 4th flip?

  Define X as number of required trails to observe the second Head (success). Thus,

$$P(X = 4) = P(\{TTHH, HTTH, THTH\}) = 3p^2\, q^2 = \binom{4-1}{2-1} p^2\, q^{4-2}.$$

Negative Binomial Distribution

A Negative Binomial experiment possesses the following two properties:

1. A Bernoulli trail with probability p of success and probability of failure $q = 1 - p$ at each step is repeated.

2. The experiment is stopped when a kth success occurs.

If X is the number of steps required until the kth success occurs, then $X \sim NB(k, p)$ with the probability function

$$b^*(x; k, p) = P(X = x) = \binom{x-1}{k-1} p^k \, q^{x-k}, \quad x = k, k+1, \ldots .$$

Expectation and Variance for Negative Binomial Distribution

If $X \sim NB(k, p)$,

- Expecation

$$\mathrm{E}(X) = \sum_{x=k}^{\infty} x \binom{x-1}{k-1} p^k \, q^{x-k} = \frac{k}{p}.$$

- Variance

$$\mathrm{Var}(X) = \mathrm{E}\left[(X - \frac{k}{p})^2 \right] = \frac{k}{p^2}.$$

Poisson Distribution-Concept

Experiments yielding numerical values of a random variable X, the number of outcomes occurring during a given time interval or in a specified region, are called Poisson experiments.

- The given time interval may be of any length, such as a minute, a day, a week, a month, or even a year. For example, a Poisson experiment can generate observations for the random variable X representing
  - the number of telephone calls received per hour by an office,
  - the number of games postponed due to rain during a baseball season.
- The specified region could be a line segment, an area, a volume, or perhaps a piece of material. In such instances, X might represent
  - the number of typing errors per page,
  - the number of accidents in an intersection.

Poisson Process

A Poisson experiment is derived from the Poisson process and possesses the following properties.

1. The number of outcomes occurring in one time interval or specified region of space is independent of the number that occur in any other disjoint time interval or region. In this sense, we say that the Poisson process has no memory.

2. The probability that a single outcome will occur during a very short time interval or in a small region is proportional to the length of the time interval or the size of the region and does not depend on the number of outcomes occurring outside this time interval or region.

3. The probability that more than one outcome will occur in such a short time interval or fall in such a small region is negligible.

Poisson Distribution

- The number X of outcomes occurring during a Poisson experiment is called a Poisson random variable, and its probability distribution is called the Poisson distribution.

- The mean number of outcomes is computed from $\mu = \lambda t$, where t is the specific "time", "distance", "area" or "volume" of interest.

- The probability function of the Poisson random variable X, representing the number of outcomes occurring in a given time interval or specified region denoted by t, is

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

where λ is the average number of outcomes per unit time, distance, area, or volume and e $=$ 2.71828. We simply write $X \sim P(\lambda t)$.

Example:

Births in a hospital occur randomly at an average rate of 2 births per hour.

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

Let X be number of births in a given hour. Then, $X \sim P(\lambda t)$ with $\lambda = 2$ and $t = 1$:

$$p(4;2) = P(X = 4) = \frac{e^{-2}\,2^4}{4!}.$$

(b) What would be the probability of observing 4 births in a given day at the hospital?

Let Y be number of births in a given day. Then, $Y \sim P(\lambda t)$ with $\lambda = 2$ and $t = 24$:

$$p(4;2\times 24) = P(Y = 4) = \frac{e^{-2\times 24}\,(2\times 24)^4}{4!}.$$

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

$$P(X \geq 4) = 1 - P(X \leq 3) = 1 - \sum_{x=0}^{3} \frac{e^{-2}\, 2^x}{x!}.$$

Table A.2 Poisson Probability Sums $\sum_{x=0}^{r} p(x;\mu)$

	μ								
r	**1.0**	**1.5**	**2.0**	**2.5**	**3.0**	**3.5**	**4.0**	**4.5**	**5.0**
0	0.3679	0.2231	0.1353	0.0821	0.0498	0.0302	0.0183	0.0111	0.0067
1	0.7358	0.5578	0.4060	0.2873	0.1991	0.1359	0.0916	0.0611	0.0404
2	0.9197	0.8088	0.6767	0.5438	0.4232	0.3208	0.2381	0.1736	0.1247
3	0.9810	0.9344	0.8571	0.7576	0.6472	0.5366	0.4335	0.3423	0.2650
4	0.9963	0.9814	0.9473	0.8912	0.8153	0.7254	0.6288	0.5321	0.4405
5	0.9994	0.9955	0.9834	0.9580	0.9161	0.8576	0.7851	0.7029	0.6160
6	0.9999	0.9991	0.9955	0.9858	0.9665	0.9347	0.8893	0.8311	0.7622
7	1.0000	0.9998	0.9989	0.9958	0.9881	0.9733	0.9489	0.9134	0.8666
8		1.0000	0.9998	0.9989	0.9962	0.9901	0.9786	0.9597	0.9319
9			1.0000	0.9997	0.9989	0.9967	0.9919	0.9829	0.9682
10				0.9999	0.9997	0.9990	0.9972	0.9933	0.9863
11				1.0000	0.9999	0.9997	0.9991	0.9976	0.9945
12					1.0000	0.9999	0.9997	0.9992	0.9980
13						1.0000	0.9999	0.9997	0.9993
14							1.0000	0.9999	0.9998
15								1.0000	0.9999
16									1.0000

Expectation and Variance for Poisson Distribution

If $X \sim P(\lambda t)$,

- Expecation

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

- Variance

$$\mathrm{Var}(X) = \mathrm{E}\left[(X - \lambda t)^2\right] = \lambda t.$$

For a proof, see Appendix A.13 of the reference book.

Continuous Uniform Distribution

- One of the simplest continuous distributions in all of statistics is the continuous uniform distribution.
- The distribution is characterized by a density function that is "flat", and thus the probability is uniform in a closed interval, say $[a, b]$.
- The density function of the continuous uniform random variable X on the interval $[a, b]$ is

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \le x \le b \\ 0 & \text{for } elsewhere. \end{cases}$$

- The mean and variance of the uniform distribution are

$$E(X) = \frac{a+b}{2}, \quad Var(X) = \frac{(b-a)^2}{2}.$$

Example:

Suppose that a large conference room at a certain company can be reserved for no more than 4 hours. Both long and short conferences occur quite often. In fact, it can be assumed that the length X of a conference has a uniform distribution on the interval $[0, 4]$.

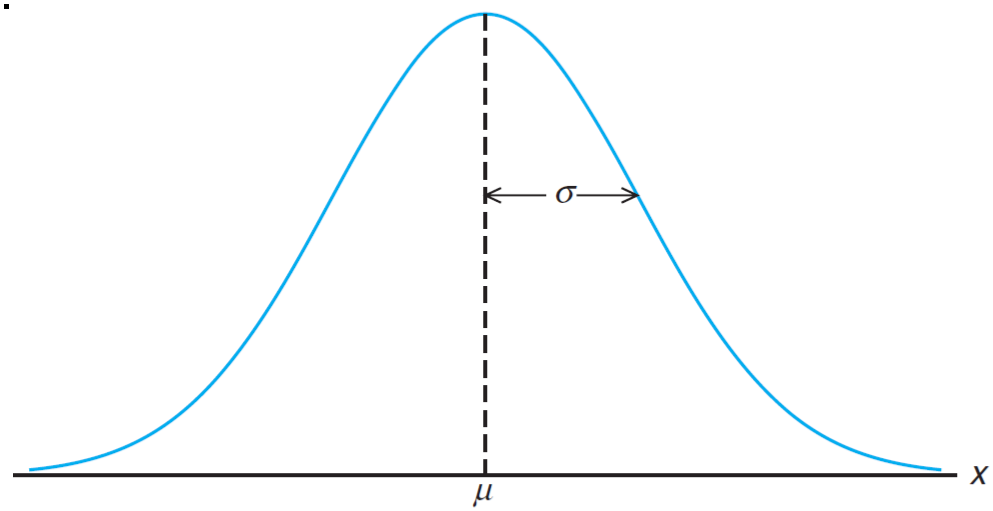
(a) What is the probability density function?

(b) What is the probability that any given conference lasts at least 3 hours?

Normal Distribution

- The most important continuous probability distribution in the entire field of statistics is the normal distribution.

- Its graph, called the normal curve, is a bell-shaped curve, which approximately describes many phenomena that occur in nature, industry, and research.



Normal Distribution

- The density of the normal random variable X, with mean μ and variance σ^2, is

$$n(x;\mu,\sigma) = f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad \mu \in \Re,\ \sigma > 0,\ x \in \Re,$$

where $\pi = 3.14159...$ and e = 2.71828....

- The normal distribution is often referred to as the Gaussian distribution

- We simply write $X \sim N(\mu, \sigma^2)$.

Properties of the Normal Curve

- The mode, which is the point on the horizontal axis where the curve is a maximum, occurs at $x = \mu$.

- The curve is symmetric about a vertical axis through the mean μ.

- The curve has its points of inflection at $x = \mu - \sigma$ and $x = \mu + \sigma$; it is concave downward if $\mu - \sigma < x < \mu + \sigma$ and is concave upward otherwise.

- The normal curve approaches the horizontal axis asymptotically as we proceed in either direction away from the mean.

- The total area under the curve and above the horizontal axis is equal to 1.

Expectation and Variance for Normal Distribution

If $X \sim N(\mu, \sigma^2)$,

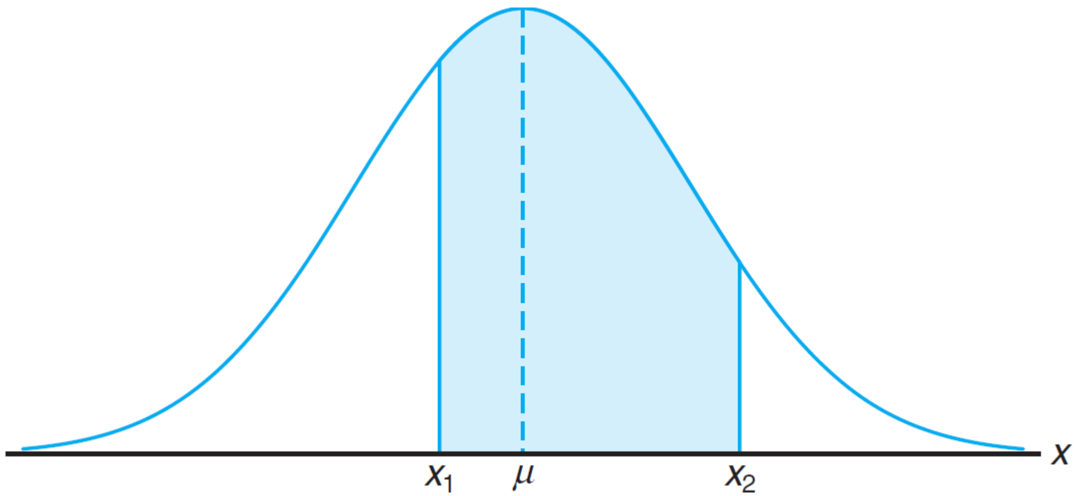
- Expecation

$$\mathrm{E}(X) = \int_{-\infty}^{+\infty} x f(x) dx = \mu.$$

- Variance

$$\mathrm{Var}(X) = \mathrm{E}\left[(X-\mu)^2\right] = \sigma^2.$$

Areas under the Normal Curve



$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x)dx = \int_{x_1}^{x_2} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx$$

Standard Normal Distribution

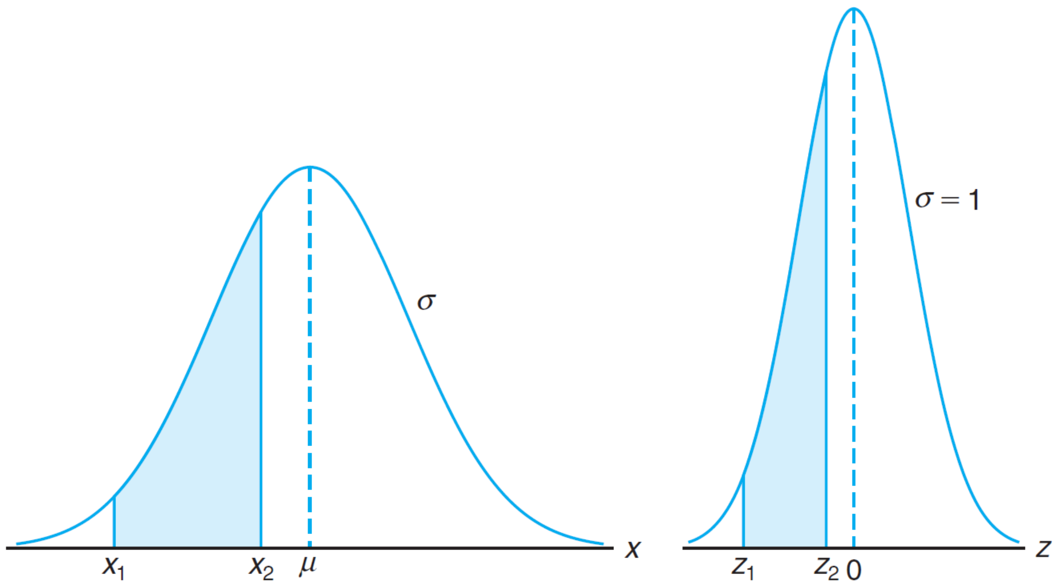
- If $X \sim N(\mu, \sigma^2)$, then $Z = \frac{X-\mu}{\sigma}$ has a standard Normal distribution, or briefly $Z \sim N(0,1)$.

- The density of the standard normal random variable Z is

$$n(z;0,1) = f(z) = \phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, \quad z \in \Re,$$

where $\pi = 3.14159...$ and e = 2.71828....

Standard Normal Distribution



$$
\begin{aligned}
P(x_1 < X < x_2) &= P\left(\frac{x_1 - \mu}{\sigma} < Z < \frac{x_2 - \mu}{\sigma}\right) \\
&= P(z_1 < Z < z_2) \\
&= \Phi(z_2) - \Phi(z_1)
\end{aligned}
$$

Example:

Given a standard normal distribution, find the area under the curve that lies to the right of $z = 0.84$.

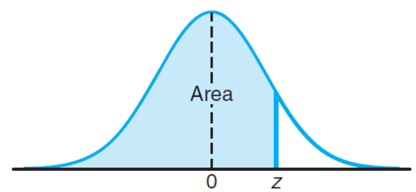


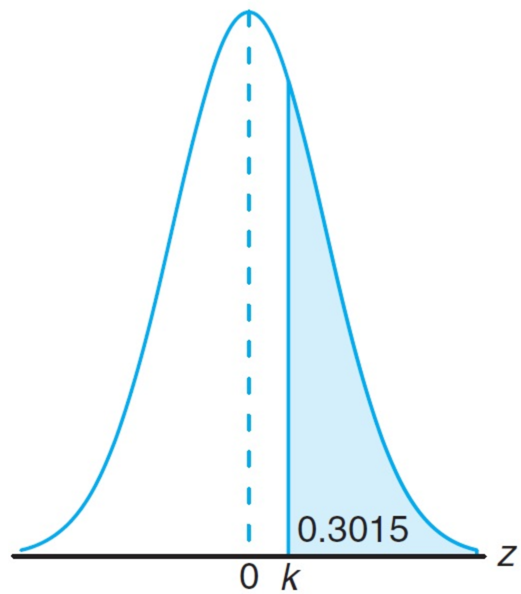
Table A.3 Areas under the Normal Curve

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389

Example:

Given a standard normal distribution, find the value of k such that

$$P(Z > k) = 0.3015.$$



Example:

An electrical firm manufactures light bulbs that have a life, before burn-out, that is normally distributed with mean equal to 300 hours and a standard deviation of 50 hours. Find the probability that a bulb burns at least 342 hours.

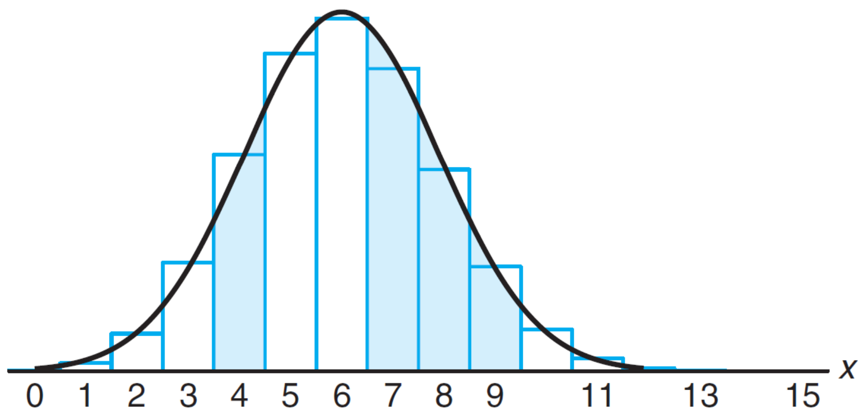
Normal Approximation to the Binomial

- If $X \sim B(n, p)$, a Binomial random variable with mean $\mu = np$ and variance $\sigma^2 = npq$, then the limiting form of the distribution of

$$Z = \frac{X - np}{\sqrt{npq}},$$

as $n \to \infty$, is the standard normal distribution $N(0, 1)$.

Example: $X \sim B(15, 0.4)$:



- Direct computation:

$$P(X = 4) = 0.1268$$

- Approximate computation:

$$\begin{aligned} P(3.5 < X < 4.5) &= P(-1.32 < Z < -0.79) \\ &= P(Z < -0.79) - P(Z \leq -1.32) \end{aligned}$$

Example:

The probability that a patient recovers from a rare blood disease is 0.4. If 100 people are known to have contracted this disease, what is the probability that fewer than 30 survive?

$$P(X < 30) = P(X \leq 29) = P\left(Z < \frac{29.5 - 100(0.4)}{\sqrt{100(0.4)(0.6)}}\right)$$

Gamma Function

- Gamma distribution is based on the Gamma function

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx, \quad \alpha > 0$$

- The Gamma function has the following properties:
  - $\Gamma(n) = (n-1)\Gamma(n-1)$, for a positive integer n;
  - $\Gamma(n) = (n-1)!$, for a positive integer n;
  - $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$, for $\alpha > 1$;
  - $\Gamma(1) = 1$;
  - $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

Gamma and Exponential Distributions

- The continuous random variable X has a Gamma Distribution, with parameters α and β, if its density function is given by

$$f(x;\alpha,\beta)=\frac{1}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\beta}},\quad \alpha>0,\ \beta>0,\ x>0.$$

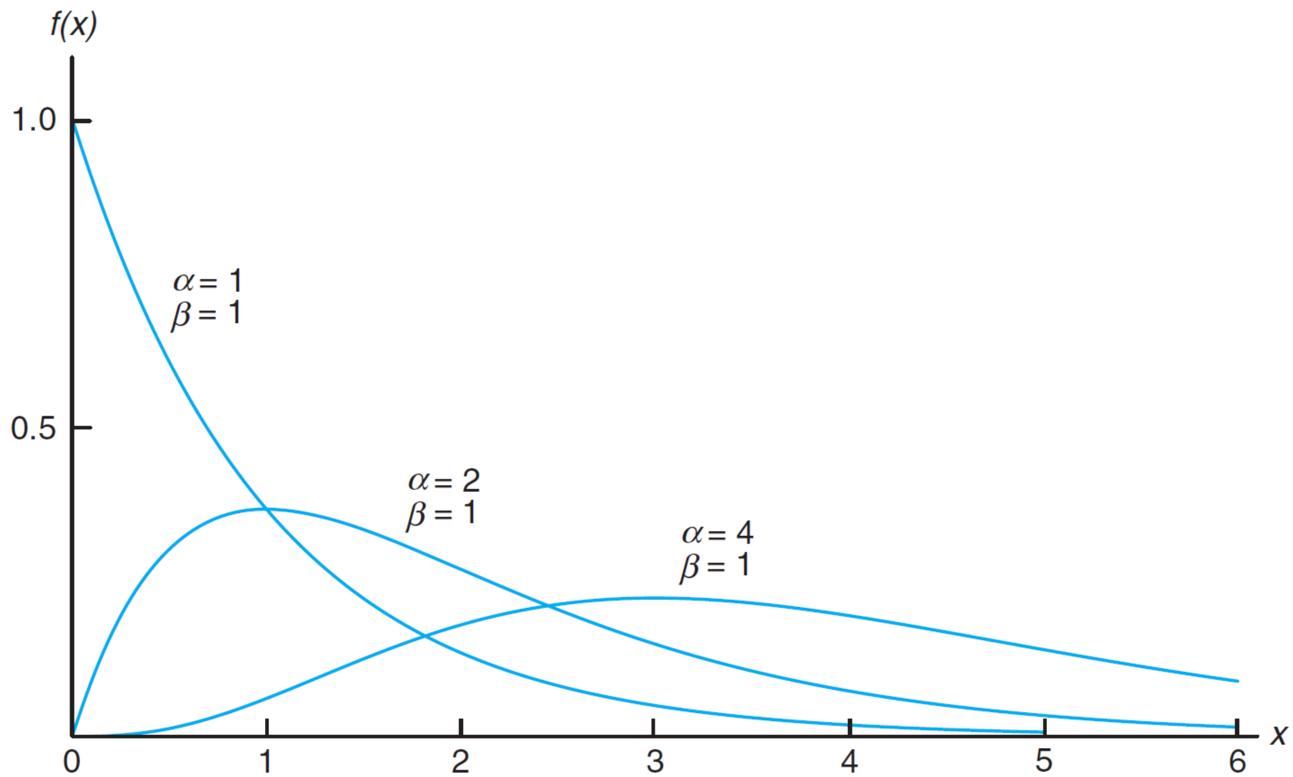
- We simply write $X \sim G(\alpha,\beta)$.

- Special Case; Exponential Distribution

$$f(x;\beta)=\frac{1}{\beta}e^{-\frac{x}{\beta}},\quad \beta>0,\ x>0.$$

- We simply write $X \sim Exp(\beta)$.

- The exponential and gamma distributions play an important role in both queuing theory and reliability problems.



Expectation and Variance for Gamma and Exponential Distributions

If $X \sim G(\alpha, \beta)$,

- Expecation $$\mathrm{E}(X) = \int_0^{+\infty} x f(x) dx = \alpha\beta.$$
- Variance $$\mathrm{Var}(X) = \mathrm{E}\left[(X - \alpha\beta)^2\right] = \alpha\beta^2.$$

If $X \sim Exp(\beta)$,

- Expecation $$\mathrm{E}(X) = \int_0^{+\infty} x f(x) dx = \beta.$$
- Variance $$\mathrm{Var}(X) = \mathrm{E}\left[(X - \beta)^2\right] = \beta^2.$$

Example:

Suppose that a system contains a certain type of component whose time, in years, to failure is given by T . The random variable T is modeled nicely by the exponential distribution with mean time to failure 5.

(a) What is the probability that a given components is still functioning at the end of 8 years?

(b) If 5 of these components are installed in different systems, what is the probability that at least 2 are still functioning at the end of 8 years?

The Memoryless Property and Its Effect on the Exponential Distribution

- If $X \sim Exp(\beta)$,

$$P(X \geq t + t_0 | X \geq t_0) = P(X \geq t).$$

Example:

Suppose that the amount of time one spends in a bank is exponentially distributed with mean 10 minutes.

- What is the probability that a customer will spend more than 15 minutes in the bank?

- What is the probability that a customer will spend more than 15 minutes in the bank given that he is still in the bank after 10 minutes?

Chi-Squared Distribution

- Another very important special case of the Gamma distribution is obtained by letting $\alpha = \frac{\nu}{2}$ and $\beta = 2$, where ν is a positive integer. The result is called the Chi-Squared Distribution.

- ν called the degrees of freedom.

- The density function is given by

$$f(x;\nu) = \frac{1}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})} x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}}, \quad \nu > 0,\, x > 0.$$

- We simply write $X \sim \chi^2_{(\nu)}$.

- $E(X) = \nu$ and $Var(X) = 2\nu$.