Chapter 5: Continuous Probability Distributions

Department of Mathematics
Izmir University of Economics

Week 7-8
2014-2015

Introduction

In this chapter we will focus on

- continuous random variables,
- cumulative distribution functions and probability density functions of continuous random variables,
- expected value, variance, and standard deviation of continuous random variables, and
- some special continuous distributions.

In the previous chapter, we developed discrete random variables and their probability distributions. Now, we extend the probability concepts to continuous random variables and their probability distributions.

Many economic and business measures such as sales, investments, consumptions, costs, and revenues can be represented by continuous random variables.

Definition:

A random variable is a *continuous random variable* if it can take any value in an interval.

Example. Consider the experiment of filling 12 ounce cans with coffee and let X be the amount of coffee in a randomly chosen can. Find the values of the random variable X.

Solution.

X can take any value between 0 and 12, that is, $0 \leq x \leq 12$ Hence, X is a continuous random variable.

Some examples of continuous random variables are:

1. The yearly income for a family.
2. The amount of oil imported into Turkey in a particular month.
3. The change in the price of a store of IBM common stock in a month.
4. The time that elapses between the installation of a new component and its failure.
5. The percentage of impurity in a batch of chemicals.

Definition:

Let X be a continuous random variable and x a specific value of the random variable X. The cumulative distribution function, $F(x)$, for a continuous random variable X expresses the probability that X does nt exceed the value of x, as a function of x. That is,

$$F(x) = P\{X \leq x\} = P\{X < x\}.$$

Note: For continuous random variables it doesn't matter whether we write "less than "or "less than or equal to "because the probability that X is precisely equal to x is 0.

Definition:

Let X be a continuous random variable with a cumulative distribution function $F(x)$, and let a and b be two possible values of X, with $a < b$. The probability that X lies between a and b is

$$P\{a < X < b\} = F(b) - F(a).$$

Example. If X is a continuous random variable with cumulative distribution function

$$F(x) = \begin{cases} 0 & \text{if } x < 0, \\ 0.001x & \text{if } 0 \leq x \leq 1000, \\ 1 & \text{if } x > 1000, \end{cases}$$

a) find $P\{X \leq 400\}$,

b) find $P\{250 < X < 750\}$.

Definition:

Let X be a continuous random variable and x any number lying in the range of that random variable. The probability density function, $f(x)$, of the random variable X is a function with the following properties:

1. $f(x) > 0$ for all values of x.
2. The area under $f(x)$ over all values of X is equal to 1, that is,

$$\int_{x} f(x)\,dx = 1.$$

3. Suppose that $f(x)$ is graphed and let a and b two possible values of the random variable X with $a < b$. Then

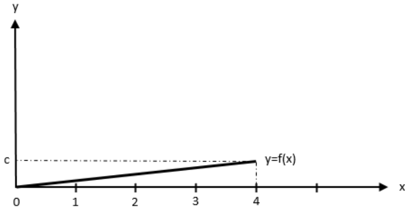
$$P\{a \leq X \leq b\} = \int_{a}^{b} f(x)\,dx.$$

4. <1-> The cumulative distribution function $F(x_0)$ is the area under $f(x)$ up to x_0. That is,

$$F(x_0) = \int_{x_m}^{x_0} f(x)\,dx,$$

where x_m is the minimum value of the random variable X.

Example. Suppose that a continuous random variable takes values on $[0, 4]$ and the graph of its probability density function is given by



a) Find c.

b) Find $P\{X \leq 3\}$.

c) Find $P\{1 \leq X \leq 2\}$.

Expectations for Continuous Random Variables

Suppose that an experiment gives results that can be represented by a continuous random variable X. If $g(X)$ is any function of X, then the expected value of $g(X)$ can be calculated by

$$E\left[g\left(X\right)\right] = \int_{x} g\left(x\right) f\left(x\right) dx.$$

Definition:

Let X be a continuous random variable. Then

- the mean of X is $\mu_X = E(X) = \int_x xf(x)\,dx$,
- the variance of X is $\sigma_X^2 = E\left((X-\mu_X)^2\right) = \int_x (x-\mu_X)^2 f(x)\,dx$,

  or alternatively, $\sigma_X^2 = E(X^2) - \mu_X^2 = \int_x x^2 f(x)\,dx - \mu_X^2$, and
- the standard deviation of X, σ_X, is the positive square root of the variance of X.

If $Y = a + bX$, where X is a continuous random variable with mean μ_X and variance σ_X^2 and a and b are constants, then

$$\mu_Y = a + b\mu_X, \quad \sigma_Y^2 = b^2\sigma_X^2, \quad \text{and} \quad \sigma_Y = |b|\sigma_X.$$

Example. If X is a continuous random variable having the probability density function

$$f(x) = \begin{cases} 0.25 & \text{if } 0 \leq x \leq 4, \\ 0 & \text{otherwise,} \end{cases}$$

a) calculate the mean and variance of X.

b) Calculate the mean, variance, and standard deviation of the random variable $Y = 2X - 3$.

The Uniform Distribution

X is a uniform random variable on the interval (a, b) if its probability density function is

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a < x < b, \\ 0 & \text{otherwise } (x \leq a \text{ or } x \geq b). \end{cases}$$

The cumulative distribution function of a uniform random variable on (a, b) is given by

$$F(x) = \begin{cases} 0 & \text{if } x \leq a, \\ \frac{x-a}{b-a} & \text{if } a < x < b, \\ 1 & \text{if } x \geq b. \end{cases}$$

The mean and variance of a uniform random variable X are

$$\mu_X = E(X) = \frac{a+b}{2} \quad \text{and} \quad \sigma_X^2 = Var(X) = \frac{(b-a)^2}{12}.$$

The Uniform Distribution

Example. If X is a continuous random variable having the probability density function

$$f(x) = \begin{cases} 0.25 & \text{if } 0 \leq x \leq 4, \\ 0 & \text{otherwise,} \end{cases}$$

a) graph the probability density function,

b) find and graph the cumulative distribution function,

c) find $P\{X < 2\}$,

d) find $P\{1 < X < 3\}$,

e) calculate the mean and variance of X.

The Uniform Distribution

Example. Let X be a continuous random variable with the probability density function

$$f(x) = \begin{cases} \frac{1}{5} & \text{if } -2 \leq x \leq 3, \\ 0 & \text{otherwise,} \end{cases}$$

a) find $P\{-2 < X < 3\}$,

b) find $P\{X > 1\}$,

c) find $P\{X \geq 1\}$.

The Uniform Distribution

Example. The incomes of all families in a particular suburb can be represented by a continuous random variable. It is known that the median income for all families in the suburb is \$60000 and that 40% of all families in the suburb have incomes above \$72000.

a) For a randomly chosen family, what is the probability that its income will be between \$60000 and \$72000?

b) If the distribution of the income is known to be uniform, what is the probability that a randomly chosen family has an income below \$65000?

The Normal Distribution

A continuous random variable X is said to have normal distribution if its probability density function is

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{\sigma^2}} \quad \text{for} \quad -\infty < x < \infty,$$

where μ and σ^2 are any numbers such that $-\infty < \mu < \infty$ and $0 < \sigma^2 < \infty$, $e = 2.71828\ldots$, and $\pi = 3.14159\ldots$.

The Normal Distribution

Properties of the Normal Distribution

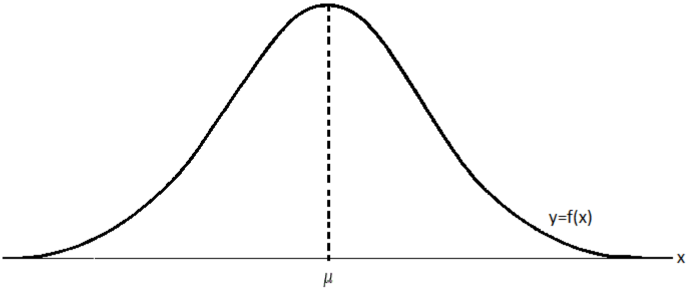
Suppose that the random variable X follows a normal distribution with parameters μ and σ^2. Then

1. $E(X) = \mu$
2. $Var(X) = \sigma^2$
3. If we know the mean and variance, we can define the normal distribution by using the notation:

$$X \sim \mathcal{N}\left(\mu, \sigma^2\right)$$

The Normal Distribution

4. <1-> The shape of the probability density function is a symmetric bell-shaped curve centered on the mean μ



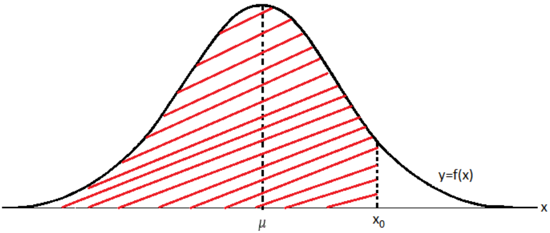
The Normal Distribution

Cumulative Distribution Function of the Normal Distribution

Suppose that $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$. Then the cumulative distribution function of X is

$$F\left(x_0\right) = P\left\{X \leq x_0\right\}.$$

This is actually the area under the normal probability density function to the left of x_0

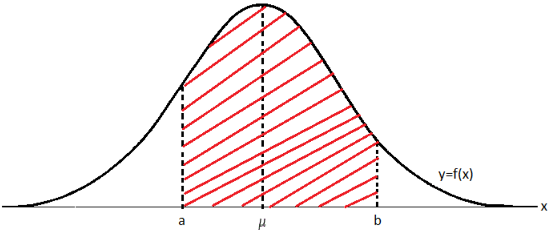


Note that, $F\left(\infty\right) = 1$.

The Normal Distribution

Let X be a normal random variable with cumulative distribution function $F(x)$ and let a and b be two possible values of the random variable X with $a < b$. Then

$$P\{a \leq X \leq b\} = F(b) - F(a).$$



The Standard Normal Distribution

Definition:

Let Z be a normal random variable with mean 0 and variance 1, that is, $Z \sim \mathcal{N}(0, 1)$. We say that Z follows a standard normal distribution. If the cumulative distribution function of Z is $F(z)$ and a and b are two possible values of Z with $a < b$, then

$$P\{a \leq Z \leq b\} = F(b) - F(a).$$

The Standard Normal Distribution

We can obtain probabilities for any normally distributed random variable by

- first converting the random variable to the standard normally distributed random variable Z using the transformation

$$Z = \frac{X - \mu}{\sigma},$$

  where $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$,
- then using the standard normal distribution table:

Cumulative Distribution of Standard Normal Distribution F(z)

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
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
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990

The Standard Normal Distribution

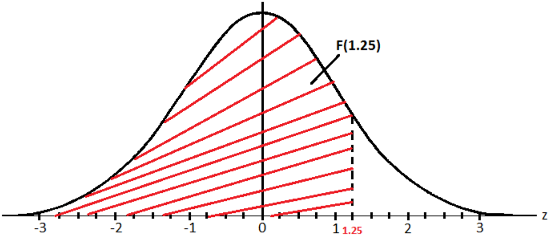
The standard normal distribution table gives the values of $F(z) = P\{Z \leq z\}$ for nonnegative values of z. For example, $F(1.25) = P\{Z \leq 1.25\} = 0.8944$

Cumulative Distribution of Standard Normal Distribution F(z)

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	[illegible]	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	[illegible]	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	[illegible]	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	[illegible]	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	[illegible]	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	[illegible]	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	[illegible]	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	[illegible]	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	[illegible]	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	[illegible]	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	[illegible]	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	[illegible]	0,8708	0,8729	[illegible]	0,8770	0,8790	0,8810	0,8830
1,2	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	[illegible]	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990

The Standard Normal Distribution

and it is actually the area of the shaded region

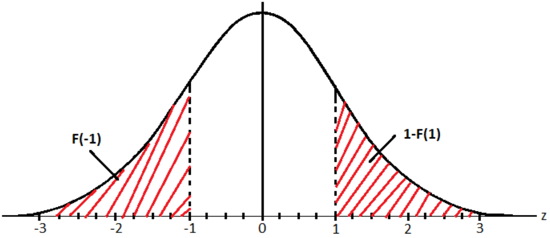


The Standard Normal Distribution

To find the cumulative probability for a negative z we use the complement of the probability of the positive value of this negative z. From symmetry, we have

$$F(-z) = P\{Z \leq -z\} = P\{Z > z\} = 1 - P\{Z \leq z\} = 1 - F(z).$$

For example, $F(-1) = 1 - F(1) = 1 - 0.8413 = 0.1587$ gives the area of



The Standard Normal Distribution

Example. Find the area under the standard normal curve that lies

a) to the left of $z = 2.5$,

b) to the left of $z = -1.55$,

c) between $z = -0.78$ and $z = 0$,

d) between $z = 0.33$ and $z = 0.66$.

The Standard Normal Distribution

Example. A random variable has a normal distribution with $\mu = 69$ and $\sigma = 5.1$. What are the probabilities that the random variable will take a value

a) less than 74.1?

b) greater than 63.9?

c) between 69 and 72.3?

d) between 66.2 and 71.8?

The Standard Normal Distribution

Example. A very large group of students obtains test scores that are normally distributed with mean 60 and standard deviation 15. Find the cutoff point for the top 10% of all students for the test scores.

The Standard Normal Distribution

Example. A random variable has a normal distribution with variance 100. Find its mean if the probability that it will take on a value less than 77.5 is 0.8264.

Normal Distribution Approximation for Binomial Distribution

Normal distribution can be used to approximate the discrete binomial distribution. This approximation allows to compute probabilities for larger sample sizes when tables are not available.

Consider a problem with n independent trials each with the probability of success p. The binomial random variable X can be written as the sum of n independent Bernoulli random variables

$$X = X_1 + X_2 + \cdots + X_n,$$

where X_i takes the value 1 if the outcome of the ith trial is "success "and 0 otherwise, with respective probabilities p and $1 - p$. Thus, X has binomial distribution with mean and variance

$$E(X) = \mu = np \quad \text{and} \quad Var(X) = \sigma^2 = np(1-p).$$

Normal Distribution Approximation for Binomial Distribution

If the number of trials (n) is large so that $np\,(1-p) > 5$, then the distribution of the random variable

$$Z = \frac{X-\mu}{\sigma} = \frac{X-np}{\sqrt{np(1-p)}}$$

is approximately a standard normal distribution.

Normal Distribution Approximation for Binomial Distribution

Example. An election forecaster has obtained a random sample of 900 voters in which 500 indicate that they will vote for Susan Chung. Assuming that there are only two candidates, should Susan Chung anticipate winning the election?

Proportion Random Variable

Definition:

A proportion random variable, P, can be computed by dividing the number of successes (X) by the sample size (n):

$$P = \frac{X}{n}$$

Then, the mean and variance of P are

$$E(P) = \mu = p \quad \text{and} \quad Var(P) = \sigma^2 = \frac{p(1-p)}{n}.$$

Proportion Random Variable

Example. An election forecaster has obtained a random sample of 900 voters in which 500 indicate that they will vote for Susan Chung. Assuming that there are only two candidates, should Susan Chung anticipate winning the election? Solve using proportion random variable.

The Exponential Distribution

The exponential distribution has been found particularly useful for waiting-line, queueing, or lifetime problems.

Definition:

The exponential random variable T ($T > 0$) has a probability density function

$$f(t) = \lambda e^{-\lambda t} \text{ for } t > 0,$$

where *lambda* > 0 is the mean number of independent arrivals or occurrences per time unit, t is the number of time units until the next arrival or occurrence, and $e = 2.71828\ldots$.

The Exponential Distribution

Let T be an exponential random variable. The cumulative distribution function of T is

$$F(t_0) = P\{T \leq t_0\} = 1 - e^{-\lambda t_0} \text{ for } t_0 > 0.$$

The mean and variance of T are

$$E(T) = \frac{1}{\lambda} \quad \text{and} \quad Var(T) = \frac{1}{\lambda^2}.$$

The Exponential Distribution

Example. Service times for customers at a library information desk can be modeled by an exponential distribution with a mean service time of 5 minutes. What is the probability that a customer service time will take longer than 10 minutes?

The Exponential Distribution

Note: The Poisson distribution provides the probability of X arrival/occurrences during a unit interval. In contrast, the exponential distribution provides the probability that an arrival/occurrence will occur during an interval of time t.

The Exponential Distribution

Example. An industrial plant in Britain with 2000 employees has a mean number of lost-time accidents per week equal to $\lambda = 0.4$ and the number of accidents follows a Poisson distribution. What is the probability that the time between accidents is less than 2 weeks?

The Exponential Distribution

Example. Times to gather preliminary information from arrivals at an outpatient clinic follow an exponential distribution with mean 15 minutes. Find the probability, for a chosen interval, that more than 18 minutes will be required.