Normal 20M

Normal

Let X follow a ***normal*** distribution with mean μ and variance σ^2, i.e.

$$X \sim \text{Normal}\left(\mu,\, \sigma^2\right)$$

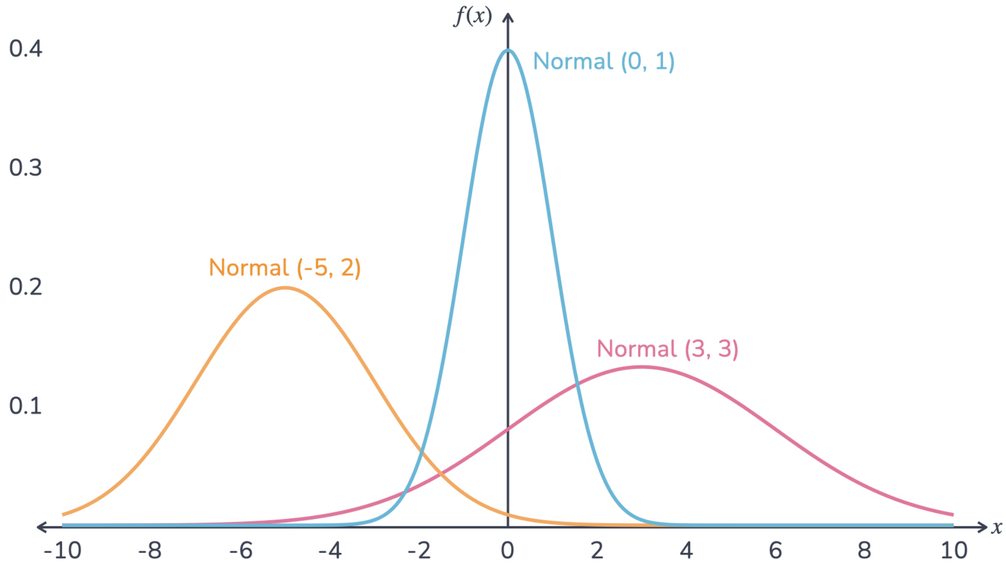
Then, X has the following PDF:

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

The normal distribution is unique in that its parameters are its mean and variance.

$$\text{E}[X] = \mu$$

$$\text{Var}[X] = \sigma^2$$



The blue curve in the middle is the PDF of a ***standard normal*** distribution, which has mean 0 and variance 1 and is usually denoted as Z.

$$Z \sim \text{Normal}\,(0, 1)$$

For a standard normal distribution, represent the CDF as $\Phi(z)$.

$$\Phi(z) = \Pr(Z \le z)$$

In addition, denote the $100p^{\text{th}}$ percentile of the standard normal distribution as z_p.

$$z_p = \Phi^{-1}(p)$$

Understanding the standard normal distribution is critical in solving all normal distribution problems, including related distributions such as the lognormal distribution.

As mentioned before, normal distribution calculators will be provided on the exam. Being familiar with these calculators is vital.

Coach's Remarks

Note that students will have access to **two** calculators.

With the **cumulative normal distribution calculator** (which we shorten to "CDF calculator" in this section), insert a percentile to get the corresponding probability.

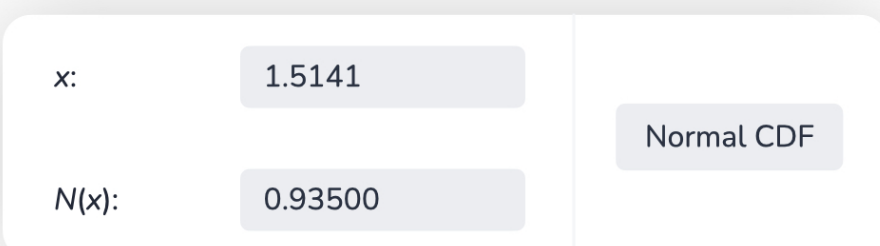
With the **inverse CDF calculator**, insert a probability to get the corresponding percentile.

Consider the following example:

Z is a standard normal random variable.

1. Calculate $\Pr(Z \leq 1.5141)$.
2. Calculate $\Pr(Z \leq -1.5141)$.

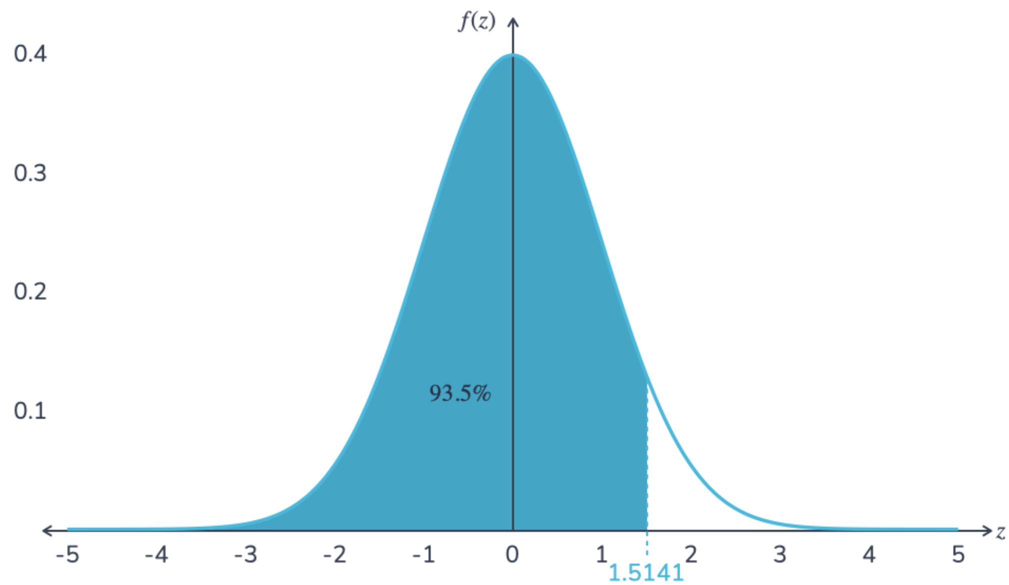
Start from the first question. Enter 1.5141 into the CDF calculator, we get



$$\Pr(Z \leq 1.5141) = \Phi(1.5141) = \mathbf{0.935}$$

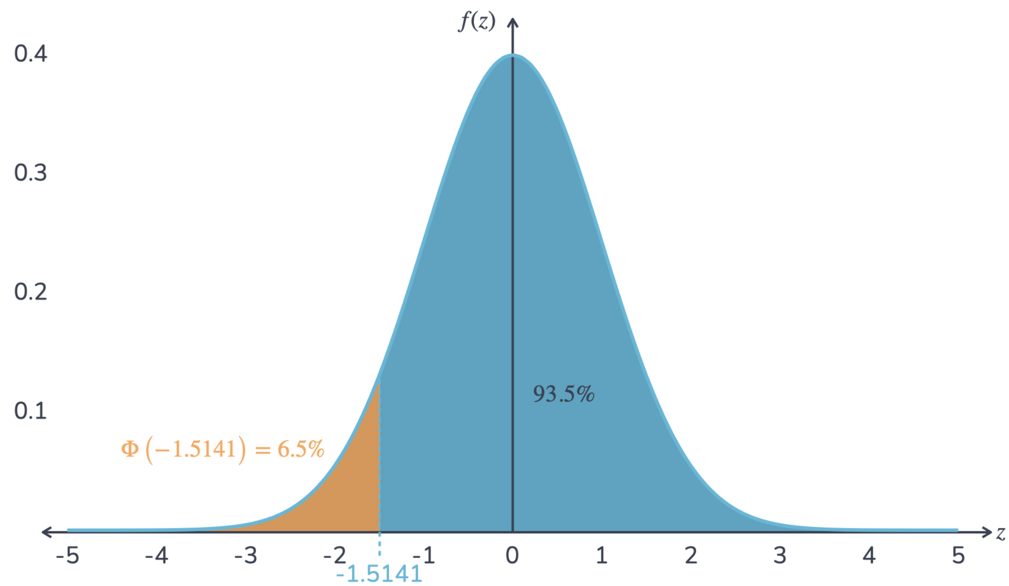
That means 1.5141 is the 93.5th percentile of the standard normal distribution, or $z_{0.935}$. Then, due to the perfect symmetry of the standard normal distribution,

$$\Pr(Z \geq -1.5141) = \Pr(Z \leq 1.5141) = 0.935$$

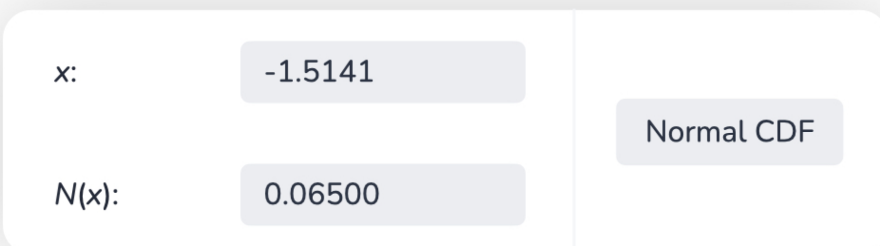


Thus, the remaining density is the answer to the second question.

$$\begin{aligned}\Pr(Z \leq -1.5141) &= \Phi(-1.5141) \\ &= 1 - \Phi(1.5141) \\ &= 1 - 0.935 \\ &= \mathbf{0.065}\end{aligned}$$



We can also get the same answer using the calculator.



$$\Pr(Z \leq -1.5141) = \Phi(-1.5141) = \mathbf{0.065}$$

The CDF of a normal distribution can be calculated by standardizing the normal random variable.

$$Z = \frac{X - \mu}{\sigma} \qquad \Leftrightarrow \qquad X = \mu + Z\sigma$$

$$\begin{aligned}\Pr(X \le x) &= \Pr(\mu + Z\sigma \le x)\\ &= \Pr\left(Z \le \frac{x-\mu}{\sigma}\right)\\ &= \Phi\left(\frac{x-\mu}{\sigma}\right)\end{aligned}$$

The same logic applies to percentiles. Let x_p be the $100p^{\text{th}}$ percentile of a normal distribution with mean μ and variance σ^2.

$$x_p = \mu + z_p\sigma$$

Consider the following example:

X follows a normal distribution with mean 3 and standard deviation 2.

1. Calculate the 97$^{\text{th}}$ percentile of X.
2. Calculate $\Pr(X \le 5)$.

Find the 97$^{\text{th}}$ standard normal percentile using the inverse CDF calculator.

N(*x*):	0.97	Inverse CDF
x:	1.88079	

$$z_{0.97} = 1.88079$$

Reverse the standardization to get the percentile for X.

$$\begin{aligned} x_{0.97} &= \mu + z_{0.97}\sigma \\ &= 3 + 1.88079\,(2) \\ &= \mathbf{6.76158} \end{aligned}$$

To calculate the probability, standardize X and use the CDF calculator.

$$\begin{aligned} \Pr(X \le 5) &= \Pr\left(\frac{X-\mu}{\sigma} \le \frac{5-3}{2}\right) \\ &= \Pr(Z \le 1) \\ &= \Phi(1) \\ &= \mathbf{0.84134} \end{aligned}$$

x:	1	Normal CDF
N(*x*):	0.84134	

Lognormal

A ***lognormal*** distribution is a log-transformed normal distribution.

For $Y \sim \text{Normal}\left(\mu,\, \sigma^2\right)$ and $X = e^Y$,

$$X \sim \text{Lognormal}\left(\mu,\, \sigma^2\right)$$

A lognormal distribution with parameters μ and σ^2 has the following PDF:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, \qquad x > 0$$

A lognormal distribution's PDF resembles a normal distribution's PDF, with the most prominent difference of $\ln x$ in the exponent in place of x.

Note that, unlike the normal distribution, the μ and σ^2 of the lognormal distribution are **not** its mean and variance. They are just parameters.

The lognormal mean and variance are as follows. The variance can be calculated from the first and second moments or using the shortcut formula below.

$$\text{E}[X] = e^{\mu + 0.5\sigma^2}$$

$$\text{Var}[X] = \text{E}[X]^2 \left(e^{\sigma^2} - 1\right)$$

The CDF of a lognormal distribution can be calculated by log-transforming the normal random variable:

$$X = e^Y \qquad \Leftrightarrow \qquad Y = \ln X$$

$$\begin{aligned}\Pr(X \le x) &= \Pr\left(e^Y \le x\right) \\ &= \Pr(Y \le \ln x) \\ &= \Pr\left(Z \le \frac{\ln x - \mu}{\sigma}\right) \\ &= \Phi\left(\frac{\ln x - \mu}{\sigma}\right)\end{aligned}$$

Note that X is the lognormal random variable, Y is the normal random variable (mean μ and variance σ^2), and Z is the standard normal random variable.

Percentiles can be calculated intuitively using what we know about normal distribution percentiles:

1. Start with the standard normal's 100p^{th} percentile.

$$z_p$$

2. Convert to the corresponding 100p^{th} percentile of a non-standard normal random variable.

$$y_p = \mu + z_p\sigma$$

3. Convert the normal percentile to its lognormal counterpart.

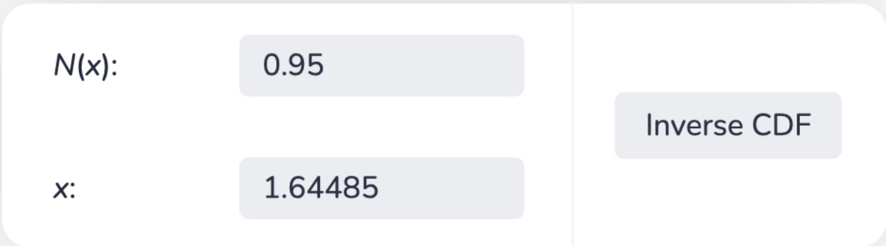
$$x_p = e^{y_p} = e^{\mu + z_p\sigma}$$

Consider the following example:

X is a lognormal random variable with parameters $\mu = 0.4$ and $\sigma^2 = 0.25$.

1. Calculate the 95$^{\text{th}}$ percentile of X.
2. Calculate $\Pr(X \leq 2)$

Find the 95$^{\text{th}}$ standard normal percentile using the inverse CDF calculator.



$$z_{0.95} = 1.64485$$

Reverse the standardization to calculate the 95$^{\text{th}}$ normal percentile.

$$y_{0.95} = 0.4 + 1.64485\sqrt{0.25} = 1.222425$$

The 95$^{\text{th}}$ lognormal percentile is

$$x_{0.95} = e^{1.222425} = \mathbf{3.39541}$$

To calculate the probability, convert the lognormal random variable to normal, and then standardize it.

$$\begin{aligned}\Pr(X \le 2) &= \Pr(\ln X \le \ln 2) \\ &= \Pr\left(\frac{\ln X - \mu}{\sigma} \le \frac{\ln 2 - 0.4}{\sqrt{0.25}}\right) \\ &= \Pr(Z \le 0.58629) \\ &= \Phi(0.58629) \\ &= \mathbf{0.72116}\end{aligned}$$

x:	0.58629	Normal CDF
N(*x*):	0.72116	

Sum of Normal Random Variables

The sum of **independent** normal random variables is a new normal random variable with mean and variance equal to the sum of the individual means and variances, respectively.

For example, assume Y_1 and Y_2 are two independent normal random variables.

$$Y_1 \sim \text{Normal}\left(\mu_1,\, \sigma_1^2\right)$$

$$Y_2 \sim \text{Normal}\left(\mu_2,\, \sigma_2^2\right)$$

Define Y as the sum.

$$Y = Y_1 + Y_2$$

Y follows a normal distribution with the following mean and variance.