Prof. Lorenzo Naranjo Spring 2025

The Lognormal Distribution

The Normal Distribution

We say that a real-valued random variable (RV) X is normally distributed with mean μ and standard deviation σ if its *probability density function* (PDF) is:

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

and we usually write $X \sim \mathcal{N}(\mu, \sigma^2)$. The parameters μ and σ are related to the first and second moments of X.

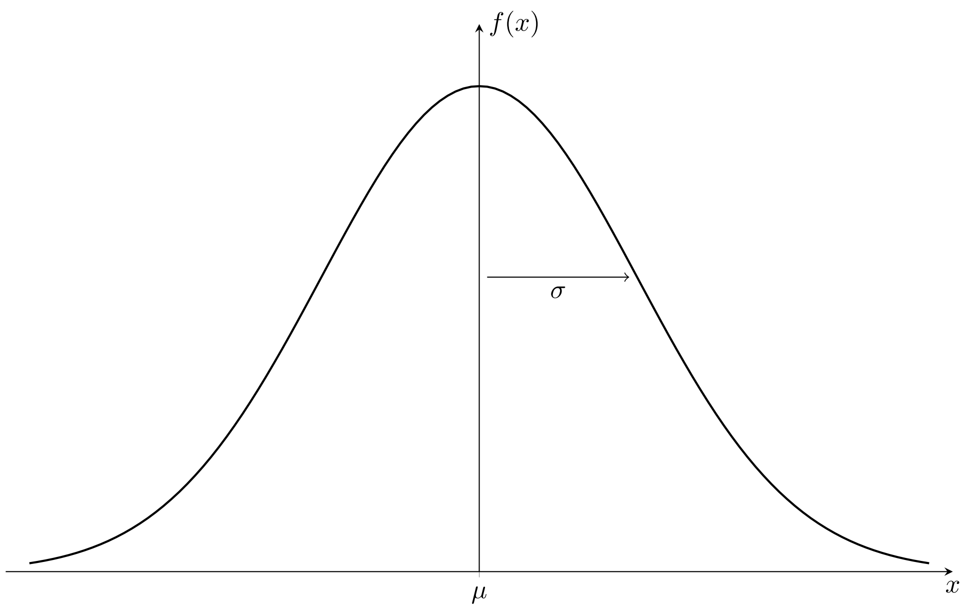

Figure 1: The figure shows the density function of a normally distributed random variable with mean μ and standard deviation σ.

Moments of the Normal Distribution

The parameter μ is the mean or expectation of X while σ denote its standard deviation. The variance of X is given by σ^2.

Proof

Let $X = \mu + \sigma Z$ where $Z \sim \mathcal{N}(0,1)$. Start by defining $f(z) = e^{-\frac{1}{2}z^2}$, which implies that $f'(z) = -ze^{-\frac{1}{2}z^2}$ and $f''(x) = z^2 e^{-\frac{1}{2}z^2} - e^{-\frac{1}{2}z^2}$. We can then write:

$$\begin{aligned} ze^{-\frac{1}{2}z^2} &= -f'(z) \\ z^2 e^{-\frac{1}{2}z^2} &= f''(x) + f(z) \end{aligned}$$

Then,

$$\begin{aligned} \mathrm{E}(Z) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} ze^{-\frac{1}{2}z^2}\, dz \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} -f'(z)\, dz \\ &= \frac{1}{\sqrt{2\pi}} \left(-f(z)\Big|_{-\infty}^{\infty} \right) \\ &= 0, \\ \mathrm{E}(Z^2) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} z^2 e^{-\frac{1}{2}z^2}\, dz \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f''(x) + f(z)\, dz \\ &= \frac{1}{\sqrt{2\pi}} \left(f'(z)\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} f(z)\, dz \right) \\ &= \frac{1}{\sqrt{2\pi}} (0 + \sqrt{2\pi}) \\ &= 1, \\ \mathrm{Var}(Z) &= \mathrm{E}(Z^2) - \mathrm{E}(Z)^2 \\ &= 1. \end{aligned}$$

Note that we used the fact that

$$\int_{-\infty}^{\infty} f(z)\,dz = \sqrt{2\pi}.$$

We can now compute $\mathsf{E}(X) = \mu + \sigma\,\mathsf{E}(Z) = \mu$ and $\mathsf{Var}(X) = \sigma^2\,\mathsf{Var}(Z) = \sigma^2$. □

As with any real-valued random variable X, in order to compute the probability that $X \leq x$ we need to integrate the density function from $-\infty$ to x:

$$\mathsf{P}(X \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(u-\mu)^2}{2\sigma^2}}\,du.$$

The function $F(x) = \mathsf{P}(X \leq x)$ is called the *cumulative distribution function* of X. The *Leibniz integral rule* implies that $F'(x) = f(x)$.

The Standard Normal Distribution

An important case of normally distributed random variables is when $\mu = 0$ and $\sigma = 1$. In this case we say that $Z \sim \mathcal{N}(0, 1)$ has the *standard normal distribution* and its cumulative distribution function is usually denoted by the capital Greek letter Φ (phi), and is defined by the integral:

$$\Phi(z) = \mathsf{P}(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx.$$

Since the integral cannot be solved in closed-form, the probability must then be obtained from a table or using a computer. For example, in R we can compute $\Phi(-0.4)$ by typing the following:

```r
pnorm(-0.4)
```

```
[1] 0.3445783
```

If you prefer to use Excel, you need to type in a cell `=norm.s.dist(-0.4,TRUE)`, which yields the same answer.

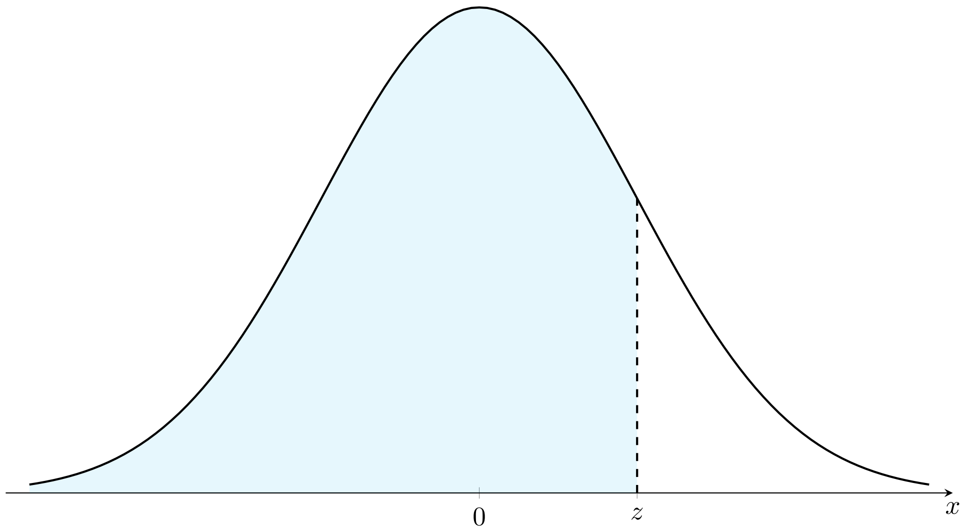

Figure 2: The blue shaded area represents $\Phi(z)$.

Left-Tail Probability

Knowing how to compute or approximate $\Phi(z)$ allows us to compute $\mathrm{P}(X \leq x)$ when $X \sim \mathcal{N}(\mu, \sigma^2)$ since $Z = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$:

$$\begin{aligned} \mathrm{P}(X \leq x) &= \mathrm{P}\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) \\ &= \mathrm{P}\left(Z \leq \frac{x-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{x-\mu}{\sigma}\right) \end{aligned}$$

where $Z = \dfrac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$ is called a Z-score.

Example 1. Suppose that $X \sim \mathcal{N}(\mu, \sigma^2)$ with $\mu = 10$ and $\sigma = 25$. What is the probability that $X \leq 0$?

$$\begin{aligned} \mathrm{P}(X \leq 0) &= \mathrm{P}\left(Z \leq \tfrac{0-10}{25}\right) \\ &= \Phi(-0.40) \\ &= 0.3446. \end{aligned}$$

□

Right-Tail Probability

For a random variable X, the *right-tail* probability is defined as $\mathrm{P}(X > x)$. Since $\mathrm{P}(X \leq x) + \mathrm{P}(X > x) = 1$, we have that:

$$\mathrm{P}(X > x) = 1 - \mathrm{P}(X \leq x).$$

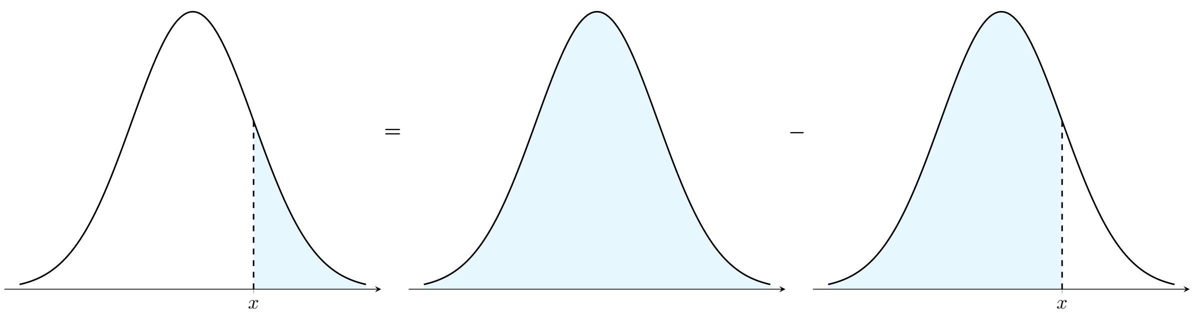

Figure 3: The right-tail probability is the probability of the whole distribution, which is one, minus the left-tail probability.

Example 2. Suppose that $X \sim \mathcal{N}(\mu, \sigma^2)$ with $\mu = 10$ and $\sigma = 25$. What is the probability that $X > 12$?

$$\begin{aligned} \mathrm{P}(X \leq 12) &= \mathrm{P}\left(Z \leq \tfrac{12-10}{25}\right) \\ &= \Phi(0.08) \\ &= 0.5319. \end{aligned}$$

Therefore, $\mathrm{P}(X > 12) = 1 - 0.5319 = 0.4681$. □

Interval Probability

The probability that a random variable X falls within an interval $(X_1, X_2]$ is given by $\mathrm{P}(x_1 < X \leq x_2) = \mathrm{P}(X \leq x_2) - \mathrm{P}(X \leq x_1)$.

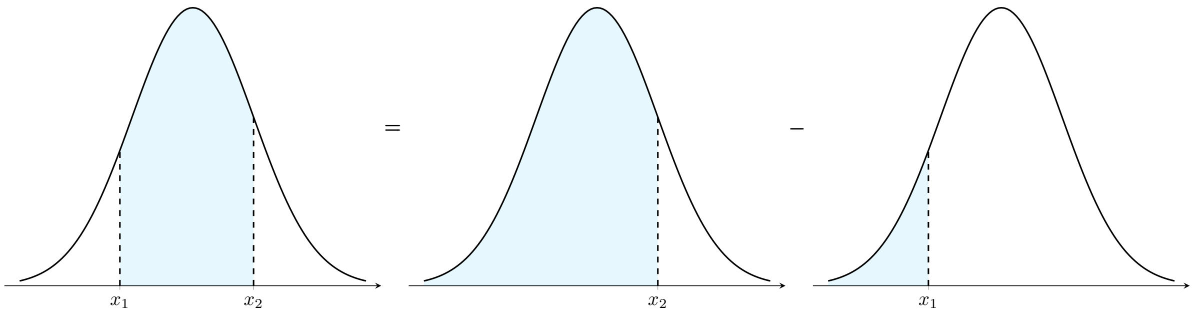

Figure 4: If you subtract the area to the left of x_1 to the area that is to the left of x_2 you obtain the probability of $x_1 < X \leq x_2$.

Example 3. Suppose that $X \sim \mathcal{N}(\mu, \sigma^2)$ with $\mu = 10$ and $\sigma = 25$. What is the probability that $2 < X \leq 14$?

$$\begin{aligned} \mathrm{P}(X \leq 14) &= \mathrm{P}\left(Z \leq \tfrac{14-10}{25}\right) \\ &= \Phi(0.16) \\ &= 0.5636, \\ \mathrm{P}(X \leq 2) &= \mathrm{P}\left(Z \leq \tfrac{2-10}{25}\right) \\ &= \Phi(-0.32) \\ &= 0.3745. \end{aligned}$$

Therefore, $\mathrm{P}(2 < X \leq 14) = 0.5636 - 0.3745 = 0.1891.$ □

The Lognormal Distribution

If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $Y = e^X$ is said to be *lognormally distributed* with the same parameters. The pdf of a lognormally distributed random variable Y can be obtained from the pdf of X.

Lognormal Density

If Y is lognormally distributed with parameters μ and σ^2, the PDF of Y is given by:

$$f(y) = \frac{1}{y\sqrt{2\pi\sigma^2}} e^{-\frac{(\ln(y)-\mu)^2}{2\sigma^2}}.$$

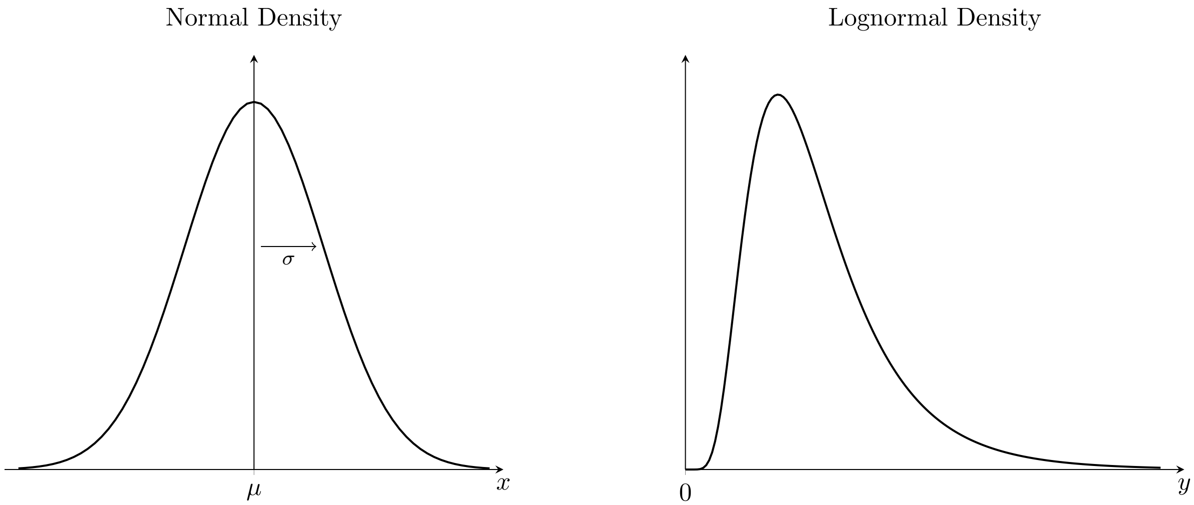

Figure 5: The figure shows the difference between a normal and a lognormal PDF with the same parameters.

Proof

Let $Y = e^X$ where $X = \mu + \sigma Z$ and $Z \sim \mathcal{N}(0,1)$. Then,

$$\begin{aligned} \mathrm{P}(Y \leq y) &= \mathrm{P}(X \leq \ln(y)) \\ &= \int_{-\infty}^{\ln(y)} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx. \end{aligned}$$

Let's define $z = e^x$. This implies that $x = \ln(z)$, which in turn implies that $dx = (1/z)dz$. Therefore,

$$\mathrm{P}(Y \leq y) = \int_{-\infty}^{y} \frac{1}{z\sqrt{2\pi\sigma^2}} e^{-\frac{(\ln(z)-\mu)^2}{2\sigma^2}}\, dz.$$

Thus, the integrand of the previous expression is the probability density function of Y. □

Unlike the normal density, the lognormal density function is not symmetric around its mean. Normally distributed variables can take values in $(-\infty, \infty)$, whereas lognormally distributed variables are always positive.

Computing Probabilities

We can use the fact that the logarithm of a lognormal random variable is normally distributed to compute cumulative probabilities.

Example 4. Let $Y = e^{4+1.5Z}$ where $Z \sim \mathcal{N}(0, 1)$. What is the probability that $Y \leq 100$?

$$\begin{aligned} \mathrm{P}(Y \leq 100) &= \mathrm{P}(e^X \leq 100) \\ &= \mathrm{P}(X \leq \ln(100)) \\ &= \mathrm{P}\left(Z \leq \tfrac{\ln(100)-4}{1.5}\right) \\ &= \Phi(0.4034) \\ &= 0.6567 \end{aligned}$$

Therefore, there is a 65.67% chance that Y is less than or equal 100. □

Moments

Moments of a Lognormal Distribution

Let $Y = e^{\mu+\sigma Z}$ where $Z \sim \mathcal{N}(0, 1)$. We have that:

$$\begin{aligned} \mathrm{E}(Y) &= e^{\mu+0.5\sigma^2} \\ \mathrm{Var}(Y) &= e^{2\mu+\sigma^2}(e^{\sigma^2} - 1) \\ \mathrm{SD}(Y) &= \mathrm{E}(Y)\sqrt{e^{\sigma^2} - 1} \end{aligned}$$

Proof

$$
\begin{aligned}
\mathrm{E}(Y) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} e^{x}\, dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}+x}\, dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-(\mu+\sigma^2))^2}{2\sigma^2}+(\mu+0.5\sigma^2)}\, dx \\
&= e^{\mu+0.5\sigma^2} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-(\mu+\sigma^2))^2}{2\sigma^2}}\, dx}_{=1} \\
&= e^{\mu+0.5\sigma^2}
\end{aligned}
$$

Using the fact that $\alpha X \sim \mathcal{N}(\alpha\mu, (\alpha\sigma)^2)$, it is also possible to compute the expectation of powers of lognormally distributed variables:

$$\mathrm{E}(Y^{\alpha}) = \mathrm{E}(e^{\alpha X}) = e^{\alpha\mu+0.5(\alpha\sigma)^2}.$$

This is useful to compute the variance and standard deviation of Y:

$$
\begin{aligned}
\mathrm{Var}(Y) &= \mathrm{E}(Y^2) - (\mathrm{E}(Y))^2 \\
&= e^{2\mu+2\sigma^2} - e^{2\mu+\sigma^2} \\
&= e^{2\mu+\sigma^2}(e^{\sigma^2} - 1) \\
\mathrm{SD}(Y) &= \sqrt{\mathrm{Var}(Y)} \\
&= \mathrm{E}(Y)\sqrt{e^{\sigma^2} - 1}
\end{aligned}
$$

□

Example 5. Let $Y = e^{4+1.5Z}$ where $Z \sim \mathcal{N}(0, 1)$. The expectation and standard deviation of Y are:

$$\mathrm{E}(Y) = e^{4+0.5(1.5^2)} = 168.17$$

$$\mathrm{SD}(Y) = 168.17\sqrt{e^{1.5^2} - 1} = 489.95$$

□

Appendix

Percentiles

For a standard normal variable Z, a *right-tail percentile* is the value z_α above which we obtain a certain probability α. Mathematically, this means finding z_α such that:

$$\mathsf{P}(Z > z_\alpha) = \alpha \Leftrightarrow \mathsf{P}(Z \leq z_\alpha) = 1 - \alpha.$$

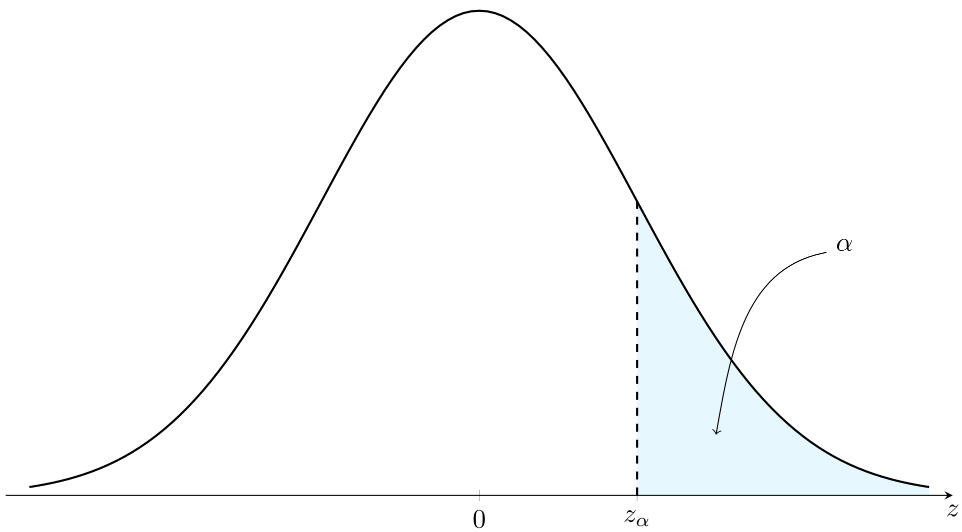

Figure 6: The right-tail percentile is the value z_α that gives an area to the right equal to α.

This implies that $\Phi(z_\alpha) = 1 - \alpha$, or $z_\alpha = \Phi^{-1}(1 - \alpha)$, where $\Phi^{-1}(\cdot)$ denotes the inverse function of $\Phi(\cdot)$. Again, there is no closed-form expression for this function and we need a computer to obtain the values. For example, say that $\alpha = 0.025$. In R we could compute $z_\alpha = \Phi^{-1}(0.975)$ by using the function `qnorm` as follows:

```
qnorm(0.975)
```

```
[1] 1.959964
```

In Excel the function `=norm.s.inv(0.975)` provides the same result.

The following table shows common values for z_α:

α	z_α
0.050	1.64
0.025	1.96
0.010	2.33
0.005	2.58

A $(1-\alpha)$ *two-sided confidence interval* (CI) defines left and right percentiles such that the probability on each side is $\alpha/2$. For a standard normal variable Z, the symmetry of its pdf implies:

$$\mathrm{P}(Z \leq -z_{\alpha/2}) = \mathrm{P}(Z > z_{\alpha/2}) = \alpha/2$$

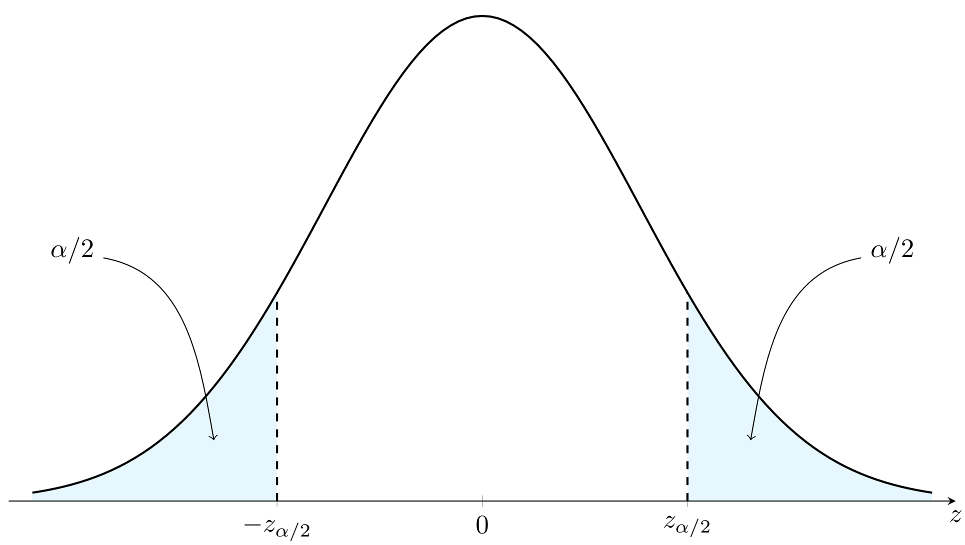

Figure 7: The areas on each side are both equal to $\alpha/2$.

Example 6. Since $z_{2.5\%} = 1.96$, the 95% confidence interval of Z is $[-1.96, 1.96]$. This means that if we randomly sample this variable 100,000 times, approximately 95,000 observations will fall inside this interval. □

If $X \sim \mathcal{N}(\mu, \sigma^2)$, its confidence interval is determined by ξ and ζ such that:

$$\begin{aligned}
&\mathrm{P}(X \le \xi) = \alpha/2 \\
&\qquad \Rightarrow \mathrm{P}(Z \le \tfrac{\xi-\mu}{\sigma}) = \alpha/2, \\
&\mathrm{P}(X > \zeta) = \alpha/2 \\
&\qquad \Rightarrow \mathrm{P}(Z > \tfrac{\zeta-\mu}{\sigma}) = \alpha/2,
\end{aligned}$$

which implies that $-z_{\alpha/2} = \frac{\xi-\mu}{\sigma}$ and $z_{\alpha/2} = \frac{\zeta-\mu}{\sigma}$.The $(1-\alpha)$ confidence interval for X is then $[\mu - z_{\alpha/2}\sigma, \mu + z_{\alpha/2}\sigma]$.

Example 7. Suppose that $X \sim \mathcal{N}(\mu, \sigma^2)$ with $\mu = 10$ and $\sigma = 25$. Since $z_{2.5\%} = 1.96$, the 95% confidence interval of X is:

$$[10 - 1.96(25), 10 + 1.96(25)] = [-39, 59].$$

□

We could also apply the same priciple for a lognormal random variable. Let $Y = e^{\mu+\sigma Z}$ where $Z \sim \mathcal{N}(0,1)$. We then have that

$$\begin{aligned}
&-z_{\alpha/2} < Z \le z_{\alpha/2} \\
&\qquad \Rightarrow \mu - \sigma z_{\alpha/2} < \mu + \sigma Z \le \mu + \sigma z_{\alpha/2} \\
&\qquad \Rightarrow e^{\mu - \sigma z_{\alpha/2}} < e^{\mu+\sigma Z} \le e^{\mu + \sigma z_{\alpha/2}}
\end{aligned}$$

The $(1-\alpha)$ confidence interval for Y (centered araound the mean of ln(Y)) is $[e^{\mu-\sigma z_{\alpha/2}}, e^{\mu+\sigma z_{\alpha/2}}]$.

Example 8. Let $Y = e^{4+1.5Z}$ where $Z \sim \mathcal{N}(0,1)$. The 95% confidence interval for Y is:

$$[e^{4-1.96(1.5)}, e^{4+1.96(1.5)}] = [2.89, 1032.71].$$

□

Partial Expectations

When pricing a call option, the payoff is positive if the option is in-the-money and zero otherwise. We usually use an indicator function to quantify this behavior:

$$\mathbb{1}_{\{Y>K\}} = \begin{cases} 0 & \text{if } Y \leq K \\ 1 & \text{if } Y > K \end{cases}$$

> Partial Expectations
>
> Let $Y = e^X$ where $X \sim \mathcal{N}(\mu, \sigma^2)$. Then we have that:
>
> $$\mathsf{E}\left(Y\mathbb{1}_{\{Y>K\}}\right) = e^{\mu+\frac{1}{2}\sigma^2}\,\Phi\left(\frac{\mu+\sigma^2-\ln(K)}{\sigma}\right)$$
> $$\mathsf{E}\left(K\mathbb{1}_{\{Y>K\}}\right) = K\,\Phi\left(\frac{\mu-\ln(K)}{\sigma}\right)$$

Proof

The first expectation can be computed as:

$$\begin{aligned}
\mathsf{E}\left(Y\mathbb{1}_{\{Y>K\}}\right) &= \int_{\ln(K)}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} e^{x}\,dx \\
&= \int_{-\infty}^{-\ln(K)} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y+\mu)^2}{2\sigma^2}} e^{-y}\,dy \\
&= \int_{-\infty}^{-\ln(K)} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y+\mu)^2}{2\sigma^2}-y}\,dy \\
&= \int_{-\infty}^{-\ln(K)} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y+(\mu+\sigma^2))^2}{2\sigma^2}+(\mu+0.5\sigma^2)}\,dy \\
&= e^{\mu+0.5\sigma^2}\int_{-\infty}^{-\ln(K)} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y+(\mu+\sigma^2))^2}{2\sigma^2}}\,dy \\
&= e^{\mu+0.5\sigma^2}\,\Phi\left(\frac{\mu+\sigma^2-\ln(K)}{\sigma}\right)
\end{aligned}$$

The second expectation yields:

$$\begin{aligned}
\mathsf{E}\left(K\mathbb{1}_{\{Y>K\}}\right) &= K\int_{\ln(K)}^{\infty}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx \\
&= K\int_{-\infty}^{-\ln(K)}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(y+\mu)^2}{2\sigma^2}}\,dy \\
&= K\,\Phi\left(\frac{\mu-\ln(K)}{\sigma}\right)
\end{aligned}$$

□

Practice Problems

Solutions to all problems can be found at lorenzonaranjo.com/fin451.

Problem 1. Suppose that X is a normally distributed random variable with mean $\mu = 12$ and standard deviation $\sigma = 20$.

a. What is the probability that $X \leq 0$?
b. What is the probability that $X \leq -4$?
c. What is the probability that $X > 8$?
d. What is the probability that $4 < X \leq 10$?

Problem 2. Suppose that X is a normally distributed random variable with mean $\mu = 10$ and standard deviation $\sigma = 20$. Compute the 90%, 95%, and 99% confidence interval for X.

Problem 3. Suppose that $X = \ln(Y)$ is a normally distributed random variable with mean $\mu = 3.9$ and standard deviation $\sigma = 15$.

a. What is the probability that $Y \leq 6$?
b. What is the probability that $Y > 4$?
c. What is the probability that $3 < Y \leq 12$?
d. What is the probability that $Y \leq 0$?

Problem 4. Suppose that X is a normally distributed variable with mean $\mu = 3.70$ and standard deviation $\sigma = 0.80$. If $Y = e^X$, what is the probability that Y is greater than 45?

Optional Practice Problems

These problems are not required to study for the exam, but can give you some good practice handling mathematical concepts discussed in the notes.

Problem 5. Suppose that $X = \ln(Y)$ is a normally distributed random variable with mean $\mu = 2.7$ and standard deviation $\sigma = 1$. Compute the 90%, 95%, and 99% confidence interval for X and report the corresponding values for Y.

Problem 6. Let $Y = e^{\mu+\sigma Z}$ where $\mu = 1, \sigma = 2$ and $Z \sim \mathcal{N}(0, 1)$. Compute:

a. $\mathrm{E}(Y)$
b. $\mathrm{SD}(Y) = \sqrt{\mathrm{E}(Y^2) - \mathrm{E}(Y)^2}$
c. $\mathrm{E}(Y^{0.3})$
d. $\mathrm{E}(Y^{-1})$