Math 3070, Applied Statistics

Section 1

September 25, 2019

Lecture Outline, 9/25

Section 4.3

- Normal Random Variables
- Mean and Variance of the Standard Normal
- Z-Transform
- CDF of a Standard Normal
- Examples

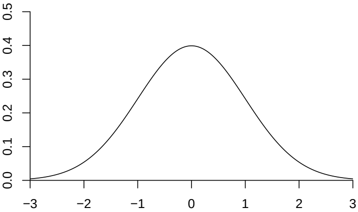
Normal Random Variables, PDF

A random variable X follows a **normal distribution** with mean μ and standard deivation σ if has the following PDF.

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

$X \sim N(\mu, \sigma)$

The $N(0,1)$ distribution is the **standard normal distribution**.



Normal Random Variables, Importance

Normal distributions often sufficiently model large sample (large n) statistics via Central Limit Theorem.

In this course, it will be used extensively when applying the central limit theorem.

Standard Normal Distribution, Integral of PDF

We want to show that the standard normal pdf is a valid pdf.
Have that $f(x) \geq 0$. Need to show $\int_{-\infty}^{\infty} f(x)dx = 1$.

We will use the special integral $\int_{-\infty}^{\infty} e^{-x^2}\ dx = \sqrt{\pi}$.
`https://en.wikipedia.org/wiki/Gaussian_integral`

Assuming this, substituting $u = x/\sqrt{2}$,

$$\begin{aligned}
\int_{-\infty}^{\infty} \phi(x)\ dx &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2}\ dx \\
&= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-u^2}\ du \\
&= \frac{1}{\sqrt{\pi}} \cdot \sqrt{\pi} = 1
\end{aligned}$$

Mean of Standard Normal

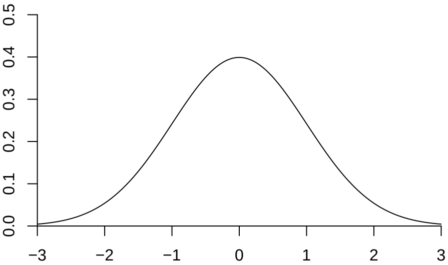
The standard normal distribution is symmetric in the sense that the pdf $\phi(x)$ is an even function, i.e., $\phi(-x) = \phi(x)$:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

Therefore the mean of a standard normal random variable is

$$E(X) = \int_{-\infty}^{\infty} x\phi(x)\ dx = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\ dx = 0$$

since the integrand is an odd function.



Variance of Standard Normal

Substituting $u = x^2/2$, note that

$$\int xe^{-x^2/2}\,dx = \int e^{-u}\,du = -e^{-u} = -e^{-x^2/2}$$

Integrating by parts,

$$\begin{aligned}
V(X) = E[(X-\mu)^2] = E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x)\,dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x \cdot xe^{-x^2/2}\,dx \\
&= \frac{1}{\sqrt{2\pi}} \left[x(-e^{-x^2/2})\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-x^2/2}\,dx \right] \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2}\,dx = 1
\end{aligned}$$

A standard normal random variable has mean 0 and variance 1.

Z-Transform, Motivation

By now, we can see that normal distributions are non-trivial to integrate. Luckily, probabilities can still be computed using Z-transforms.

Given $Z \sim N(0,1)$ or Z is a standard normal,

$$\sigma Z + \mu \sim N(\mu, \sigma).$$

The equivalently or more commonly used fact is the **Z-Transform**:

$$X \sim N(\mu, \sigma) \to \frac{X - \mu}{\sigma} \sim N(0,1)$$

Use the Z-Transform to change integrals or probabilities involving and normal random variable into ones involving the standard normal. The later can be looked up in tables such as Table A3.

Z-Transform, Proof and Calculation

Let $X \sim N(\mu, \sigma)$ and $Z \sim N(0,1)$. If the PDF of X and $\sigma Z + \mu$ match, then they follow the same distribution.

Start with the CDF of $\sigma Z + \mu$.

$$\begin{aligned} P(\sigma Z + \mu < x) &= P\left(Z < \frac{x-\mu}{\sigma}\right) \\ &= \int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-y^2}{2}\right) dy \end{aligned}$$

Differentiate to show that the PDFs are the same.

$$\begin{aligned} \frac{d}{dx} P(\sigma Z + \mu < x) &= \frac{d}{dx} \int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-y^2}{2}\right) dy \\ &= \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-(x-u)^2}{2\sigma^2}\right) \frac{1}{\sigma} \end{aligned}$$

Note: Fundamental Theorem of Calculus and Chain Rule used.

Z-Transform, Proof and Calculation

The derivative of the CDF of $\sigma Z + \mu$ and PDF of X are equal meaning that they behave the exact same way as random variables. They not the same variable, but have their events have the same probabilities.

$$P(X < x) = P\left(\frac{X-\mu}{\sigma} < \frac{x-\mu}{\sigma}\right) = P\left(Z < \frac{x-\mu}{\sigma}\right)$$

Moreover, this gives us a way to directly compute probabilities many events involving any random variable in terms of standard random variables. Notice that the left hand side is difficult if not impossible to integrate while the right hand side can be looked up as long as all non-random variables are known.

Z-Transform, Mean and Variance of Normal Distributions

Since X and $\sigma Z + \mu$ have the same PDF.

$$E[X] = E[\sigma Z + \mu] = \sigma E[Z] + \mu = \mu$$

$$V(X) = V(\sigma Z + \mu) = \sigma^2 V(Z) = \sigma^2$$

$$\text{standard deviation of } X = \sigma$$

CDF of a Standard Normal

The CDF of the a standard normal random variable $N(0,1)$ is denoted below. It is denoted since it frequently appears in calculations involving normal random variables.

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(\frac{-y^2}{2}\right) dy$$

Values of $\Phi(x)$ can be looked up by using Table A.3. Your calculator should have the ability to look up values as well.

CDF of a Standard Normal, Critical Values

The critical value z_α is defined as

$$\alpha = 1 - \Phi(z_\alpha) = P(Z \geq z_\alpha)$$

and is equal to the $100(1-\alpha)^{th}$ percentile.
Due to symmetry of the standard normal,

$$\alpha = \Phi(-z_\alpha) = P(Z \leq -z_\alpha)$$

$68 - 95 - 99.7$ Rule:

- $P(-1 < Z < 1) = P(-\sigma < X - \mu < \sigma) \approx 68\%$
- $P(-2 < Z < 2) = P(-2\sigma < X - \mu < 2\sigma) \approx 95\%$
- $P(-3 < Z < 3) = P(-3\sigma < X - \mu < 3\sigma) \approx 99.7\%$

Simple Example, Using Z-Transforms

The weight of an Altoid X follows a normal distribution with a mean of 16 grams and standard deviation of 1.2 grams. What is the probability that a randomly selected Altoid less than 17 grams in weight?

$$\begin{aligned} P(X < 17) &= P\left(\frac{X - 16}{1.2} < \frac{17 - 16}{1.2}\right) \\ &= \Phi\left(\frac{17 - 16}{1.2}\right) \\ &\approx 0.7967 \end{aligned}$$

Example, Using Z-Transforms with Absolute Values

The weight of an Altoid X follows a normal distribution with a mean of 16 grams and standard deviation of 1.2 grams. What is the probability that a randomly selected Altoid is no more than 0.5 grams away from 17 grams in weight?

$$\begin{aligned}
P(|X-17|<0.5) &= P(-0.5 < X - 17 < 0.5) \\
&= P(16.5 < X < 17.5) \\
&= P\left(\frac{16.5-16}{1.2} < \frac{X-16}{1.2} < \frac{17.5-16}{1.2}\right) \\
&= P\left(\frac{16.5-16}{1.2} < Z < \frac{17.5-16}{1.2}\right) \\
&= \Phi\left(\frac{17.5-16}{1.2}\right) - \Phi\left(\frac{16.5-16}{1.2}\right) \\
&\approx 0.8944 - 0.6628 = 0.2316
\end{aligned}$$

Example, Percentiles

The weight of an Altoid X follows a normal distribution with a mean of 16 grams and standard deviation of 1.2 grams. Under what weight are 75% of Altoids?

Find z_{75}

$$P(0.68) \approx 0.75 \to z_{75} \approx 0.68$$

Use Z-Transform.

$$\begin{aligned} 0.75 \approx P(Z < 0.68) &= P(1.2Z + 16 < 1.2 \cdot 0.68 + 16) \\ &= P(X < 16.816) \end{aligned}$$

Roughly 75% of Altoids are under 16.816 grams.

Example, Percentiles with Absolute Values

X follows a normal distribution with a mean of 4 and standard deviation of 0.4. Determine c so that

$$P(c \leq |X - 4|) = 0.5.$$

$$\begin{aligned}
0.5 = P(c \leq |X-4|) &= 1 - P(|X-4| < c) = 1 - P(-c < X - 4 < c) \\
&= 1 - P\left(-\frac{c}{0.4} < \frac{X-4}{0.4} < \frac{c}{0.4}\right) \\
&= 1 - P\left(-\frac{c}{0.4} < Z < \frac{c}{0.4}\right) \\
&= 1 - \left[\Phi\left(\frac{c}{0.4}\right) - \Phi\left(-\frac{c}{0.4}\right)\right] \\
&= 1 - \left[\Phi\left(\frac{c}{0.4}\right) - \left(1 - \Phi\left(\frac{c}{0.4}\right)\right)\right] \\
&= 2 - 2\Phi\left(\frac{c}{0.4}\right)
\end{aligned}$$

Example, Percentiles with Absolute Values

X follows a normal distribution with a mean of 4 and standard deviation of 0.4. Determine c so that

$$P(c \leq |X - 4|) = 0.5.$$

$$0.25 = 1 - \Phi\left(\frac{c}{0.4}\right) \to \Phi\left(\frac{c}{0.4}\right) = 0.75$$

Find the critical value, $z_{75} \approx 0.68$

$$\frac{c}{0.4} \approx 0.68 \to c \approx 0.272$$

Summary

- $X \sim N(\mu, \sigma)$ means X is normal distributed with mean μ and standard deviation σ. Its PDF is

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

- $X \sim N(0, 1)$ is a standard normal random variable.
- Z-Transform

$$\frac{X-\mu}{\sigma} \sim N(0,1) \qquad \sigma Z + \mu \sim N(\mu, \sigma)$$

- CDF of Z is denoted as $\Phi(x)$, useful in calculations.
- $\alpha = 1 - \Phi(z_\alpha)$ and $\alpha = \Phi(-z_\alpha)$, the distribution is symmetric.
- $68 - 95 - 99.7$ Rule
- For any random variable, $P(|X| < c) = P(-c < X < c)$.