AMES Class Notes – Week 7, Monday: Integrals and Probability

Andie Creel

9^{th} October, 2023

1 Probability density function

Consider the random variable Y. Random variables aren't random and they aren't variables. They're really more like functions.

1.1 Normal Distribution

If Y is distributed normally it will look like

$$Y = f(y) = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{y-\mu}{\sigma})^2} dy = 1 \tag{1}$$

the inside of the integral is the probability density function. The integral of a pdf must equal 1.

1.2 Example

Consider the random variable

$$Y = f(y) = 2ye^{-y^2} \tag{2}$$

and the support is $[0, \infty)$. Our probability density function (pdf) is $f(y)$.
Our cumulative distribution function (CDF) is

$$\begin{aligned} CDF &= \int_0^{\infty} 2ye^{-y^2} dy && (3) \\ &= 2\int_0^{\infty} ye^{-y^2} dy && (4) \end{aligned}$$

Let $u = -y^2$, $du = -2ydy$

$$\begin{aligned} &= 2\int_0^{\infty} ye^u \frac{du}{-2y} && (5) \\ &= -1\int_0^{\infty} e^u du && (6) \\ &= -1e^u \Big|_0^{\infty} && (7) \\ &= -1e^{-y^2} \Big|_0^{\infty} && (8) \\ &= 0 - -1 && (9) \\ &= 1 && (10) \end{aligned}$$

Which integrates to 1 after doing a u substitution, and so we know we have a proper pdf.

What if you want to find the **median**? Set the CDF equal to 0.5 because the median is the 50th percentile.

The unknown is the point on the x-axis where the area under the pdf is equal to 0.5 (which is defined as our median). Start with equation 8 from our previous integral derivation because we know the integral is equal to that. But replace the upper bound to an unknown $\tilde{y}$ and set equal to 0.5.

$$\begin{align}
-1e^{-y^2}\Big|_0^{\tilde{y}} &= 0.5 \tag{11}\\
-e^{-\tilde{y}^2} - -1 &= 0.5 \tag{12}\\
-e^{-\tilde{y}^2} &= -0.5 \tag{13}\\
e^{-\tilde{y}^2} &= 0.5 \tag{14}\\
ln(e^{-\tilde{y}^2}) &= ln(1/2) \tag{15}\\
-\tilde{y}^2 &= ln(1/2) \tag{16}\\
\tilde{y}^2 &= -ln(1/2) \tag{17}\\
\tilde{y} &= \sqrt{-ln(1/2)} \tag{18}
\end{align}$$

so we know that the median value is $\sqrt{-ln(1/2)}$.

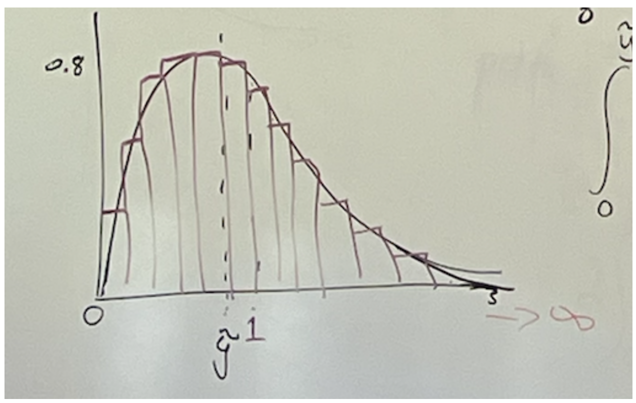

Figure 1: Histograph

1.3 Mean

To find a mean, we sum the values and divide it by the number of people we summed over. This is the same as multiplying everyone's value by $\frac{1}{N}$.

$$\begin{align}
E[Y] = \bar{y} &= \frac{1}{N}\sum_n y_n \tag{19}\\
&= \sum_n \frac{1}{N} y_n \tag{20}
\end{align}$$

We can use the pdf to get our mean! The pdf would replace the $\frac{1}{N}$ and the sum would be replaced by the integral.

$$E[Y] = \bar{y} = \int_0^\infty 2ye^{-y^2} * ydy \tag{21}$$

Important: I let the pdf be $f(y)$ for the mean and variance examples.

$$f(y) = pdf(y) \tag{22}$$

The general rule for finding the mean of a random variable by using the pdf $f(y)$ of that random variable is

$$E[Y] = \int_{\Omega} y f(y) dy \tag{23}$$

where Ω is the support of y.

1.4 Function of the random variable

Now, let's say you're interest in the function of a mean. You're interested in $g(y)$ rather that y itself.

$$E[g(y)] = \int_{\Omega} g(y) f(y) dy \tag{24}$$

where $f(y)$ continues to be the pdf of Y. This doesn't break jensen's inequality.

1.5 Variance

We an use the pdf to find the variance of Y, as well.

$$VAR(Y) = \int_{\Omega} y^2 f(y) dy - E(y)^2 \tag{25}$$

2 Uniform Distribution Example

Consider the uniform ditribution's pdf

$$f(y) = \frac{1}{\theta_2 - \theta_1} \tag{26}$$

Let $\theta_2 = 1$ and $\theta_1 = 0$

$$f(y) = \frac{1}{1-0} \tag{27}$$

$$f(y) = 1 \tag{28}$$

Prove $f(y)$ **is a pdf** by showing it integrates to 1.

$$\int_0^1 1 dy = y\Big|_0^1 = 1 - 0 = 1 \tag{29}$$

What's the mean?

$$E(Y) = \int_0^1 y * 1 dy = \frac{1}{2} y^2 \Big|_0^1 = \frac{1}{2} \tag{30}$$

What's the median?

$$Med(Y) = \int_0^{\tilde{y}} dy = 0.5 \tag{31}$$

$$y\Big|_0^{\tilde{y}} = 0.5 \tag{32}$$

$$\tilde{y} - 0 = 0.5 \tag{33}$$

What's the variance?

$$
\begin{aligned}
Var(Y) &= \int_0^1 y^2 dy - \frac{1}{2}^2 && (34)\\
&= \frac{1}{3}y^3\Big|_0^1 - \frac{1}{2}^2 && (35)\\
&= \frac{1}{3} - \frac{1}{4} && (36)\\
&= \frac{1}{12} && (37)
\end{aligned}
$$