AP STATISTICS
TOPIC VI: CONTINUOUS PROBABILITY SPACES

PAUL L. BAILEY

1. Sigma Algebras

Definition 1. Let S be a set and let $\mathcal{E} \subset \mathcal{P}(S)$.

We say that $\mathcal{E}$ is a *sigma algebra* on S if

- **(S1)** $\mathcal{E}$ is nonempty;
- **(S2)** $A^c \in \mathcal{E}$ for any $A \in \mathcal{E}$.
- **(S3)** $\cup_{i=1}^{\infty} A_i \in \mathcal{E}$ for any infinite sequence $\{A_i \mid i \in \mathbb{N}\} \subset \mathcal{E}$;

That is, a sigma algebra on S is a nonempty collection of subsets of S which is closed under complement and countably infinite unions.

Proposition 1. *Let S be a set and let $\mathcal{E}$ be a sigma algebra on S. Then*

- **(a)** $\varnothing, S \in \mathcal{E}$;
- **(b)** $A \cup B \in \mathcal{E}$ *for any* $A, B \in \mathcal{E}$;
- **(c)** $A \cap B \in \mathcal{E}$ *for any* $A, B \in \mathcal{E}$;
- **(d)** $\cup_{i=1}^{n} A_i \in \mathcal{E}$ *for any finite sequence* $\{A_1, \dots, A_n\} \subset \mathcal{E}$;
- **(e)** $\cap_{i=1}^{n} A_i \in \mathcal{E}$ *for any finite sequence* $\{A_1, \dots, A_n\} \subset \mathcal{E}$;
- **(f)** $\cap_{i=1}^{\infty} A_i \in \mathcal{E}$ *for any infinite sequence* $\{A_i \mid i \in \mathbb{N}\} \subset \mathcal{E}$.

Thus a sigma algebra on S is a collection of subsets of S which contains the empty set and the whole set, and is closed under complement, countable unions, and countable intersections.

The intersection of sigma algebras is a sigma algebra; thus given any collection $\mathcal{C}$ of subsets of S, the intersection of all sigma algebras containing $\mathcal{C}$ is a (sigma) algebra on S; it is the smallest sigma algebra on S which contains $\mathcal{C}$, and is known as the *sigma algebra generated by* $\mathcal{C}$.

The (sigma) algebra generated by $\mathcal{C}$ can be explicitly obtained by taking all possible unions of sets in $\mathcal{C}$, then taking their complements, then taking unions of these and the complements, and continuing the process indefinitely.

Date: September 19, 2023.

2. Probability Spaces

Definition 2. A *probability space* $(S, \mathcal{E}, P)$ is a set S, called a *sample space*, together with a sigma algebra $\mathcal{E}$ of subsets of S, and a function

$$P : \mathcal{E} \to [0, 1]$$

satisfying

- **(P1)** $P(S) = 1$;
- **(P2)** $P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ for any sequence $\{A_i \mid i \in \mathbb{N}\} \subset \mathcal{E}$ with $A_i \cap A_j = \varnothing$ when $i \neq j$.

The elements of S are called *outcomes*. The members of $\mathcal{E}$ are called *events*. The function P is called a *probability measure*. The number $P(E)$ is called *probability* of event E.

Proposition 2. *Let $(S, \mathcal{E}, P)$ be a probability space. Let $A, B \in \mathcal{E}$. Then*

- **(a)** $P(\varnothing) = 0$*;*
- **(b)** $P(A^c) = 1 - P(A)$*;*
- **(c)** $A \subset B \Rightarrow P(A) \leq P(B)$*;*
- **(d)** $P(A \cup B) = P(A) + P(B) - P(A \cap B)$*;*

Corollary 1. Boole's Inequality
Let $(S, \mathcal{E}, P)$ be a probability space. Let $A, B \in \mathcal{E}$. Then

$$P(A \cup B) \leq P(A) + P(B).$$

3. Continuous Random Variables

Definition 3. Let $(S, \mathcal{E}, P)$ be a probability space and let $X : S \to \mathbb{R}$ be a function. We say that X is a *continuous random variable* if $X^{-1}((-\infty, b]) \in \mathcal{E}$ for every $b \in\in \mathbb{R}$. Define $P(X \in [a, b]) = P(X^{-1}([a, b]))$.

A function $\rho : \mathbb{R} \to \mathbb{R}$ is called a *probability density function* for X if

$$P(a \leq X \leq b) = \int_a^b \rho(x)\, dx.$$

The *mean* of X is

$$\mu = \int_{-\infty}^{\infty} x\rho(x)\, dx.$$

The *standard deviation* of X is

$$\sigma = \sqrt{\int_{-\infty}^{\infty} (x - \mu)^2 \rho(x)\, dx}.$$

4. Normal Distribution

4.1. **The Derivation.** Consider a function ρ with these properties:

(a) Its rate of growth at x is proportional to $x\rho(x)$;
(b) ρ has a local maximum at $x = 0$;
(c) ρ has an inflection point at $x = 1$;
(d) the area under the curve equals 1.

By **(a)**, ρ satisfies the differential equation

$$\rho'(x) = Bx\rho(x)$$

for some constant B. Separate the variables to solve this differential equation:

$$\int \frac{\rho'(x)}{\rho(x)}\,dx = \int Bx\,dx,$$

so

$$\log(\rho(x)) = \frac{Bx^2}{2} + C,$$

and letting $A = e^C$, we have

$$\rho(x) = Ae^{\frac{Bx^2}{2}},$$

where A is positive.

By **(b)**, ρ has a local maximum at $x = 0$. Now

$$\rho'(x) = ABxe^{\frac{Bx^2}{2}},$$

which equals zero only if $x = 0$; for this to indicate a local maximum, we must have $ABx < 0$ for $x > 0$; this indicates that B is negative.

By **(c)**, ρ has an inflection point at $x = 1$. Now

$$\rho''(x) = ABe^{\frac{Bx^2}{2}}[1 + Bx^2],$$

which equals zero when $1 + Bx^2 = 0$, or $x = \pm\frac{1}{\sqrt{-B}}$. So we must have $B = -1$. Thus

$$\rho(x) = Ae^{\frac{-x^2}{2}}.$$

By **(d)**, $\int_{\mathbb{R}} \rho(x)\,dx = 1$. Using vector calculus, it is possible to compute that

$$\int_{\mathbb{R}} e^{\frac{-x^2}{2}}\,dx = \sqrt{2\pi}.$$

Thus $A = \frac{1}{\sqrt{2\pi}}$, so

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

4.2. **The Limit of the Binomial Distributions.** We wish to see how to view the normal distribution as the limit of binomial distributions. In order to compare the binomial distributions for increasing values of n, we wish to shift and compress the distribution so that the mean is zero and the variance is one.

Let X_n denote a random variable with range $\{0, h, 2h, \ldots, nh\}$ with probabilities given by successive terms of $(p+q)^n$. We have

$$P(X_n = rh) = \frac{n!}{r!(n-r!)}.$$

Then the mean is $\mu = nph$ and the variance is $\sigma^2 = npqh^2$.

Our initial goal is to understand the rate of change of the binomial distribution. Let $y = P(X = rh)$ and $y' = P(X = (r+1)h)$. Set $\Delta y = y' - y$, so that

$$\begin{aligned}
\Delta y &= \frac{n!}{(r+1)!(n-r-1)}p^{r+1}q^{n-r-1} - \frac{n!}{r!(n-r)!}p^r q^{n-r} \\
&= \frac{n!(n-r)p - n!(r+1)q}{(r+1)!(n-r)!}p^r q^{n-r-1} \\
&= \frac{n!p^r q^{n-r-1}}{(r+1)!(n-r)!}\Big((n-r)p - (r+1)q\Big) \\
&= \frac{n!p^r q^{n-r-1}}{(r+1)!(n-r)!}\Big(np - r(p+q) - q\Big) \\
&= \frac{n!p^r q^{n-r}}{r!(n-r)!} \cdot \frac{1}{(r+1)q} \cdot \Big(np - r - q\Big) \\
&= y \cdot \frac{np - r - q}{(r+1)q} \quad \text{(Equation 1).}
\end{aligned}$$

Let $x = rh - \mu = rh - nph$, so that x is now measured from the mean. Now

$$r = \frac{x}{h} + np \quad \text{and} \quad r + 1 = \frac{x}{h} + np + 1.$$

Thus

$$(np - r)h = -x \quad \text{and} \quad (r+1)h = x + h + nph.$$

Multiply top and bottom of the righthand side of Equation 1 by h^2 to get

$$\begin{aligned}
\Delta y &= y \cdot \frac{\Big((np-r)h - qh\Big)h}{\Big(x + h + nph\Big)qh} \\
&= y \cdot \frac{(-x - qh)h}{npqh^2 + (x+h)qh} \quad \text{(Equation 2).}
\end{aligned}$$

Recall that the variance is $\sigma^2 = npqh^2$. We now set $h = (npq)^{-1/2}$, which fixes $\sigma = 1$. We note that $h \to 0$ as $n \to \infty$. Set $\Delta x = h$ and divide both sides of the last equation by h to get

$$\frac{\Delta y}{\Delta x} = y \cdot \frac{(-x - hq)}{1 + (x+h)qh}.$$

Take the limit as $n \to \infty$ to compute that derivative:

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{h \to 0} y \cdot \frac{(-x - hq)}{1 + (x+h)qh} = -yx.$$

4.3. **The Differential Equation.** Thus the limit of the binomial distributions, normalized so that $\mu = 0$ and $\sigma^2 = 1$, satisfies the differential equation

$$\frac{dy}{dx} = -yx.$$

We solve this by separation of the variables:

$$\begin{aligned}
\frac{1}{y}\frac{dy}{dx} &= -x \\
\int \frac{1}{y}\frac{dy}{dx}\,dx &= \int -x\,dx \\
\int \frac{1}{y}\,dy &= \int -x\,dx \\
\log y &= -\frac{x^2}{2} + c \\
y &= e^{-x^2/2+C}
\end{aligned}$$

Letting $A = e^C$, we have

$$y = Ae^{-x^2/2}.$$

4.4. **The Constant of Integration.** We now show how a change of variable allows us to compute $\int_{\mathbb{R}} e^{-x^2/2}\,dx$, without actually finding an antiderivative for e^{-x^2}.

Proposition 3.

$$\int_{\mathbb{R}} e^{-x^2}\,dx = \sqrt{\pi}.$$

Proof. Let $I = \int_{\mathbb{R}} e^{-x^2}\,dx$. Then

$$\begin{aligned}
I^2 &= \left(\int_{\mathbb{R}} e^{-x^2}\,dx\right)\left(\int_{\mathbb{R}} e^{-x^2}\,dx\right) \\
&= \left(\int_{\mathbb{R}} e^{-x^2}\,dx\right)\left(\int_{\mathbb{R}} e^{-y^2}\,dy\right) && \text{since } x \text{ is a dummy variable} \\
&= \left(\int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-x^2}\,dx\right) e^{-y^2}\,dy\right) && \text{since } I \text{ is a constant} \\
&= \iint_{\mathbb{R}^2} e^{-x^2} e^{-y^2}\,dx\,dy && \text{since } y \text{ is constant with respect to } x \\
&= \iint_{\mathbb{R}^2} e^{-(x^2+y^2)}\,dx\,dy && \text{by properties of exponents}
\end{aligned}$$

We now use a change of variables, coverting rectangular to polar coordinates. The Jacobian for this transformation is $dx\,dy = r\,dr\,d\theta$. Since $x^2 + y^2 = r^2$, we have

$$\begin{aligned}
I^2 &= \iint_{\mathbb{R}^2} e^{-(x^2+y^2)}\,dx\,dy \\
&= \int_0^{2\pi}\int_0^{\infty} re^{-r^2}\,dr\,d\theta && \text{now let } u = -r^2, \text{ so } du = -\frac{1}{2}rdr \\
&= -\frac{1}{2}\int_0^{2\pi} \left[e^{-r^2}\right]_0^{\infty} d\theta \\
&= -\frac{1}{2}\int_0^{2\pi} [0-1]\,d\theta \\
&= \frac{1}{2}\int_0^{2\pi} d\theta \\
&= \pi.
\end{aligned}$$

Thus $I = \sqrt{\pi}$. □

Corollary 2.

$$\int_{\mathbb{R}} e^{-x^2/2}\,dx = \sqrt{2\pi}.$$

Proof. Let $u = \frac{x}{\sqrt{2}}$. Thus $du = \frac{dx}{\sqrt{2}}$, so $dx = \sqrt{2}du$, so

$$\int_{\mathbb{R}} e^{-x^2/2}\,dx = \sqrt{2}\int_{\mathbb{R}} e^{-u^2}\,du = \sqrt{2}\sqrt{\pi} = \sqrt{2\pi}.$$

□

Definition 4. Let $(S, \mathcal{E}, P)$ be a probability space and let $X : S \to \mathbb{R}$ be a random variable. We say the X is *normally distributed* if it has probability density function

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

In this case, the mean of X is μ and the standard deviation of X is σ.

The *standard normal distribution* is a normal distribution with $\mu = 0$ and $\sigma = 1$:

$$\rho : \mathbb{R} \to \mathbb{R} \quad \text{given by} \quad \rho(x) = \frac{1}{\sqrt{2\pi e^{x^2}}}.$$

Let X be a normally distributed random variable with distribution function ρ_X. The *z-value* of $x \in \mathbb{R}$ with respect to X is the signed number of standard deviations that x is away from the mean μ. This is clearly given by the following function.

The *z-transform* is the function

$$z : \mathbb{R} \to \mathbb{R} \quad \text{given by} \quad z(x) = \frac{x - \mu}{\sigma}.$$

Given a z-value, we may retrieve the original *raw score* by noting that $x = \sigma z + \mu$.

Let $Z = \frac{X-\mu}{\sigma}$. Then $P(a \leq X \leq b) = P(z(a) \leq Z \leq z(b))$. We use the to compute probabilities for any normal distribution using the standard normal distribution.

Example 1. At a large hospital, the average blood sodium value is $\mu = 137$ with standard deviation $\sigma = 3.94$. Assume that the sodium values are normally distributed.

(a) Find the probability of a sodium value in the range $140 < x < 145$.
(b) Find the probability of a sodium value in the range $x > 150$.

Solution. For **(a)**, we have

$$\begin{aligned} P(140 < X < 145) &= P(z(140) < Z < z(145)) \\ &= P\left(\frac{140-137}{3.94} < Z < \frac{145-137}{3.94}\right) \\ &= P(0.761 < Z < 2.03) \\ &= P(Z < 2.03) - P(Z < 0.761) \\ &= .9788 - .7764 \\ &= .2024. \end{aligned}$$

So, the probability is approximately 20 %.

For **(b)**, we have

$$\begin{aligned} P(150 < X) &= P(z(150) < Z) \\ &= P\left(\frac{150-137}{3.94} < Z\right) \\ &= P(3.30 < Z) \\ &= 1 - P(Z < 3.30) \\ &= 1 - .996 \\ &= .004. \end{aligned}$$

So, in this case, the probability is about .4 %, or about 1 in 250 patients. □

Normal Approximation to the Binomial Distribution
Use $\mu = np$ and $\sigma = \sqrt{npq}$.

Example 2. About 53 percent of voters in 2008 voted for Obama. Given a random sample of 892 voters, find the probability that more than 500 of them voted for Obama.

Solution. Using the binomial distribution, we would set $p = .53$ and $n = 892$, compute $P(X = r) = \binom{n}{r}p^r(1-p)^{n-r}$ for each $r > 500$, and add these numbers.

Using the approximation, set $\mu = np = 892(.53) = 472.76$ and $\sigma = \sqrt{npq} = \sqrt{892(.53)(.47)} = 14.9$. Then

$$P(X > 500) = P\left(Z > \frac{500-472.76}{14.9}\right) = 1 - P(Z < 1.83) = 1 - .9664 = .0336.$$

□

Normal Approximation to the Poisson Distribution
Use $\mu = \lambda$ and $\sigma = \sqrt{\lambda}$.

Department of Mathematics, Paragon Science Academy
Email address: `pbailey@sonoranschools.org`