Chapter 1

MathStat Preliminaries

1.1

Wait. Where are we going?

Exponential pdf

If $X \sim \mathrm{Exponential}(\lambda)$, then X has the pdf $f(x) = \lambda \exp(-\lambda x) I_{[0,\infty]}(x)$.

Exponential pdf visualized

Sample

A sample of n realization of X, denoted $X_1, X_2, \ldots, X_n$, can be used to approximate the distribution of X.

Histogram for 10 realizations

A histogram of 10 realizations from an Exponential(15.2).

Histogram for 1000 realizations

A histogram of 1000 realizations from an Exponential(15.2).

1.1.1

A very special trick

Approach

We can avoid doing actual integration if we can manipulate the integrand to look more like a pdf, which must integrate to 1 of the range of the random variable.

Example: Determine

$$\int_0^\infty 3\exp(-2x)\,dx.$$

Special trick (cont)

1.2

Transformations of Random Variables

1.2.1

The discrete case and the binomial distribution

The binomial pmf

$X \sim \text{Binomial}(n, p)$ has the pmf

$$f(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} I_{\{0,1,...,n\}}(x).$$

Determine the pmf of $Y := n - X$.

1.2.2

The Continuous Case and the Gamma Distribution

Continous Transformation PDF

Let X be a continuous random variable with pdf $f_X(x)$. Let $Y = g(X)$, where g is invertible and differentiable. Then the pdf for Y is

$$f_Y(y) = f_X\left(g^{-1}(y)\right)\left|\frac{d}{dy}g^{-1}(y)\right|.$$

Proof

Proof (cont)

Proof (cont)

Proof (cont)

Example 1.2.2

A continuous random variable $X \sim \mathrm{Gamma}(\alpha, \beta)$ has pdf
$f_X(x) = \dfrac{1}{\Gamma(\alpha)}\beta^\alpha x^{\alpha-1}\exp(-\beta x)I_{(0,\infty)}(x).$

Example 1.2.2 (cont)

Let $Y := cX$ with $c > 0$. Determine the pdf of Y.

Example 1.2.2 (cont)

The shape parameter

The α term is the “shape” parameter of the Gamma distribution.
The kernel (non-constant) part of the pdf of the Gamma distribution is $x^{\alpha-1}\exp(-\beta x)I_{(0,\infty)}(x)$.

The shape parameter

The exponential function part is an unscaled exponential density.

- $\exp(-\beta x)$ dominates $x^{\alpha-1}$ for large x.
- For small x, the $x^{\alpha-1}$ term controls the shape.

Shape parameter visualized

The α terms controls the shape of the Gamma distribution.

The Gamma Function

The pdf for the gamma distribution is defined using the **gamma function**, denoted by $\Gamma(\alpha)$.

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} \exp(-x) \, dx.$$

The Gamma Function (cont)

The Gamma Function (cont)

The Gamma Function (cont)

1.3

Bivariate Transformations

Bivariate Transformation PDF

Suppose the X_1 and X_2 are jointly continuous random variables with pdf $f_{X_1,X_2}(x_1, x_2)$. Let $Y_1 = g_1(X_1, X_2)$ and $Y_2 = g_2(X_1, X_2)$. For h_1 and h_2 differentiable, let

$X_1 = h_1(Y_1, Y_2) \text{ and } X_2 = h_2(Y_1, Y_2).$

Then

$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}\left(h_1(y_1, y_2), h_2(y_1, y_2)\right)|J|,$

where $|J|$ is the absolute value of the Jacobian.

Bivariate transformation pdf continued

The absolute value of the Jacobian is

$$J = \det \begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{bmatrix}.$$

Example 1.3.1

Let $X_1, X_2 \overset{i.i.d.}{\sim} \mathrm{Gamma}(\alpha, \beta)$
Determine the pdf of
$Y = \dfrac{X_1}{X_1 + X_2}$.

Example 1.3.1 (cont)

Example 1.3.1 (cont)

Example 1.3.1 (cont)

Example 1.3.1 (cont)

The Beta Distribution

The continuous random variable $X \sim \mathrm{Beta}(a, b)$ had pdf

$$f(x) = \frac{1}{\mathcal{B}(a, b)} x^{a-1}(1-x)^{b-1} I_{(0,1)}(x),$$

for $a, b > 0$, where

$$\mathcal{B}(a, b) = \int_0^1 x^{a-1}(1-x)^{b-1}\, dx$$

denotes the **beta function**, which normalizes the kernel of the Beta distribution.

The Beta Distribution (cont)

The Beta distribution is a flexible distribution for modeling a random variable between 0 and 1.
Note: The Uniform(0, 1) distribution is a special case of the Beta distribution with

The Beta Distribution (cont)

The Beta Function

$$\mathcal{B}(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a + b)}.$$

Proof:

The Beta Function (cont)

1.4

Minimums and Maximums

Order statistics notation

Let $X_1, X_2, \ldots, X_n$ denote a sample of observations. The order statistics are denoted as:

- $X_{(1)} = \min(X_1, X_2, \ldots, X_n)$.
- $X_{(2)} = \text{2nd smallest of } X_1, X_2, \ldots, X_n$.
- $X_{(n)} = \max(X_1, X_2, \ldots X_n)$.

1.4.2

The distribution of a minimum by example

Example 1.4.1

Let $X_1, X_2, \ldots, X_n \overset{i.i.d.}{\sim} \text{Exponential}(\lambda)$. Determine the distribution of $X_{(1)}$.

Example 1.4.1 (cont)

Example 1.4.1 (cont)

Example 1.4.1 (cont)

Example 1.4.1 (cont)

1.4.3

The general distribution of minimums and maximums

Distribution of the minimum

Let $X_1, X_2, \ldots, X_n$ be a random sample from a continuous distribution. Determine the distribution of $X_{(1)}$.

Distribution of the minimum (cont)

Distribution of the minimum (cont)

Distribution of the minimum (cont)

PDF for a minimum

Let $X_1, X_2, \ldots, X_n$ be a random sample from a continuous distribution with pdf f and cdf X.
The pdf for $X_{(1)}$ is
$f_{X_{(1)}} = n[1 - F(x)]^{n-1} f(x)$.

Distribution of the maximum

Let $X_1, X_2, \ldots, X_n$ be a random sample from a continuous distribution. Determine the distribution of $X_{(n)}$.

Distribution of the maximum (cont)

Distribution of the maximum (cont)

Distribution of the maximum (cont)

PDF for a maximum

Let $X_1, X_2, \ldots, X_n$ be a random sample from a continuous distribution with pdf f and cdf X.
The pdf for $X_{(n)}$ is
$f_{X_{(n)}} = n[F(x)]^{n-1} f(x).$

1.5

Moment Generating Functions (mgfs)

Definition of the mgf

The moment generating function (mgf) of a random variable X is $M_X(t) = E[e^{tX}]$.

Uses of mgfs

Mgfs are often useful in two settings:

1. Determining the distribution of a transformation of 1 or more random variables (particularly the sum of n i.i.d. random variables.)
2. Determine the “moments” of a random variable, i.e., $E(X), E(X^2), E(X^3), \ldots$.

1.5.1

The expectation of a function X

Example 1.5.1

Suppose the random variable X has the following distribution.

x	**-1**	**0**	**1**
$f(x)$	4/12	3/12	5/12

Determine $E(X^2)$.

Example 1.5.1 (cont)

Example 1.5.1 (cont)

Law of the Unconscious Statistician

If X is a discrete random variable, then

$$E[g(X)] = \sum_x g(x) f_X(x),$$

where $f_X(x)$ is the pmf of X.

If X is a continuous random variable, then

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx,$$

where $f_X(x)$ is the pdf of X.

Proof (specific case)

Proof (specific case)

Proof (specific case)

Proof (specific case)

1.5.2

Finding mgfs and the Poisson distribution

The Poisson distribution

A discrete random variable $X \sim \mathrm{Poisson}(\lambda)$ if it has the pmf

$$f_X(x) = \frac{e^{-\lambda}\lambda^x}{x!} I_{\{0,1,2,\ldots\}}(x).$$

Poisson interpretation

A Poisson random variable can be thought of as the number of arrivals observed in a fixed amount of time where:

1. The mean time between arrivals is constant.
2. The number of arrivals in non-overlapping time periods are independent.

Example 1.5.2

Determine the mgf for the Poisson distribution

Example 1.5.2 (cont)

Example 1.5.2 (cont)

Example 1.5.3

Suppose that $X \sim \mathrm{Exponential}(\lambda)$. Determine the mgf of X.

Example 1.5.3 (cont)

Example 1.5.3 (cont)

Example 1.5.4

Suppose that $X \sim N(\mu, \sigma^2)$ with pdf

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

Determine the mgf of X.

Example 1.5.4 (cont)

Example 1.5.4 (cont)

Example 1.5.4 (cont)

1.5.3

Finding moments

Defining moments

The **moments** of a random variable X are expectations of the form $E(X^k)$ for $k = 1, 2, 3, \ldots$.

Computing moments

The kth moment of a random variable X can be computed as $M_X^{(k)}(0)$, where $M_X^{(k)}(0)$ is the kth derivative of the mgf of X evaluated at zero.

Mgf moment proof

Mgf moment proof (cont)

Mgf moment proof (cont)

Example 1.5.5

Determine the mean and variance of the $\mathrm{Poisson}(\lambda)$ distribution.

Example 1.5.5 (cont)

Example 1.5.5 (cont)

1.5.4

The mgf uniquely determines a distribution

The DNA of the mgf

The mgf of a random variable uniquely determines its distribution.
Example: If we find the mgf of a random variable X is $\exp[15(e^t - 1)]$, then $X \sim \text{Poisson}(15)$ since it matches the mgf of a Poisson random variable with $\lambda = 15$.

1.5.6

Sums of iid random variables

Mgf of sum of independent r.v.s

Let $X_1, X_2, \ldots, X_n$ be independent random variables. The mgf of

$$Y := \sum_{i=1}^{n} X_i$$

is

$$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(t).$$

Proof

Proof (cont)

Proof (cont)

Mgf of sum of i.i.d. r.v.s

Let $X_1, X_2, \ldots, X_n \stackrel{i.i.d.}{\sim} F$. The mgf of

$$Y := \sum_{i=1}^{n} X_i$$

is

$$M_Y(t) = [M_X(t)]^n.$$

Example 1.5.6

Suppose that $X_1, X_2, \ldots, X_n \overset{i.i.d.}{\sim} \mathrm{Poisson}(\lambda)$. Determine the distribution of $Y := \sum_{i=1}^{n} X_i$.

Example 1.5.6 (cont)

Example 1.5.6 (cont)

Example 1.5.7

Suppose that $X_1, X_2, \ldots, X_n \overset{i.i.d.}{\sim} \text{Exponential}(\lambda)$. Determine the distribution of $Y := \sum_{i=1}^{n} X_i$.

Example 1.5.7 (cont)

Example 1.5.7 (cont)

Example 1.5.8

Suppose that $X_1, X_2, \dots, X_n$ be independent random variables with $X_i \sim \mathrm{Poisson}(\lambda_i)$. Determine the distribution of $Y := \sum_{i=1}^{n} X_i$.

Example 1.5.8 (cont)

Example 1.5.8 (cont)

Example 1.5.8 (cont)

1.6

General Order Statistics

1.6.2

The joint distribution of the minimum and maximum

Min/Max cdf

For an i.i.d. sample of continuous random variables $X_1, X_2, \ldots, X_n$ with cdf F, the joint cdf of $(X_{(1)}, X_{(n)})$ is
$F_{X_{(1)},X_{(n)}}(x, y) = [F(y)]^n - [F(y) - F(x)]^n,$
for $x < y$ and
$F_{X_{(1)},X_{(n)}}(x, y) = [F(y)]^n$
for $x \geq y$.

Proof of min/max cdf

Proof of min/max cdf (cont)

Proof of min/max cdf (cont)

Min/Max pdf

For an i.i.d. sample of continuous random variables $X_1, X_2, \ldots, X_n$ with cdf F, the joint pdf of $(X_{(1)}, X_{(n)})$ is
$f_{X_{(1)}, X_{(n)}}(x, y) = n(n-1)[[F(y) - F(x)]^{n-2} f(x) f(y),$
for $x < y$.

Proof of min/max pdf

Example 1.6.1

Determine the joint pdf of $(X_{(1)}, X_{(15)})$ when $X_1, X_2, \ldots, X_{15} \overset{i.i.d.}{\sim} \text{Uniform}(0, 1)$.

Example 1.6.1 (cont)

1.6.3

The distribution of all order statistics

PDF of all order statistics

For an i.i.d. sample of continuous random variables $X_1, X_2, \ldots, X_n$ with pdf f, the pdf of all n order statistics is

$$f_{X_{(1)},X_{(2)},\ldots,X_{(n)}}(x_1, x_2, \ldots, x_n)$$
$$= n! f(x_1) f(x_2) \cdots f(x_n) I(x_1 < x_2 < \cdots < x_{n-1} < x_n).$$

For $y_1 < x_1 \le y_2 < x_2 < \cdots \le y_n < x_n$

$$P(y_1 < X_{(1)} \le x_1, y_2 < X_{(2)} \le x_2, \ldots, y_n < X_{(n)} \le x_n)$$
$$= P(\{y_1 < X_1 \le x_1, y_2 < X_2 \le x_2, \ldots, y_n < X_n \le x_n\} \cup$$
$$\{y_1 < X_2 \le x_1, y_2 < X_1 \le x_2, y_3 < X_3 \le x_3, \ldots, y_n < X_n \le x_n\} \cup$$
$$\vdots$$

$n!$ permutations

Proof $= n!\, P(y_1 < X_1 \le x_1,\ y_2 < X_2 \le x_2, \ldots, y_n < X_n \le x_n)$

$$\overset{\text{ind.}}{=} n!\, P(y_1 < X_1 \le x_1) P(y_2 < X_2 \le x_2) \cdots P(y_n < X_n \le x_n)$$

$$= n! \prod_{i=1}^{n} P(y_i < X_i \le x_i)$$

$$= n! \prod_{i=1}^{n} [F(x_i) - F(y_i)]$$

A simplified version of the first fundamental theorem of calc. is

For $x \in [a,b]$, $\frac{d}{dx} \int_a^x f(t)\, dt = f(x)$.

$$P(y_1 < X_{(1)} \le x_1, \ldots, y_n < X_{(n)} \le x_n) = \int_{y_n}^{x_n} \cdots \int_{y_1}^{x_1} f_{X_{(1)}, \ldots, X_{(n)}}(u_1, \ldots, u_n)\, du_1 \cdots du_n \Leftarrow$$

$$= n! \prod_{i=1}^{n} [F(x_i) - F(y_i)] \Leftarrow$$

$$\frac{d}{dx_n} \cdots \frac{d}{dx_1} P(y_1 < X_{(1)} \le x_1, \ldots, y_n < X_{(n)} \le x_n) = \frac{d}{dx_n} \cdots \frac{d}{dx_1} \int_{y_n}^{x_n} \cdots \int_{y_1}^{x_1} f_{X_{(1)}, \ldots, X_{(n)}}(u_1, \ldots, u_n)\, du_1 \cdots du_n$$

$$= \frac{d}{dx_n} \cdots \frac{d}{dx_1} n! \prod_{i=1}^{n} [F(x_i) - F(y_i)]$$

Proof

Claim: $\frac{d}{dx_n} \cdots \frac{d}{dx_1} \underbrace{n! \prod_{i=1}^{n} [F(x_i) - F(y_i)]} = n! f(x_1) \cdots f(x_n)$

For $n=1$, $\frac{d}{dx_1} \{ \underline{1!} [F(x_1) - \underline{F(\cancel{y_1})}] \} = 1! f(x_1)$ ✓

For $n=2$, $\frac{d}{dx_2} \frac{d}{dx_1} \{ \underline{2! [F(x_1) - F(\cancel{y_1})] \underbrace{[F(x_2) - F(y_2)]}} \}$

$\frac{d}{dx_2} \underline{2! f(x_1)} [F(x_2) - F(\cancel{y_1})]$

$= 2! f(x_1) f(x_2)$ ✓

Assume true for $n=k-1$, show for $n=k$

For $n=k$,

$\frac{d}{dx_n} \cdots \frac{d}{dx_1} n! \prod_{i=1}^{n} [F(x_i) - F(y_i)]$

$= \frac{d}{dx_n} \cdots \frac{d}{dx_1} \underbrace{n [F(x_n) - F(y_n)]} (n-1)! \prod_{i=1}^{n-1} [F(x_i) - F(y_i)]$

$$= \frac{d}{dx_n} n\left[F(x_n) - F(y_n)\right] \underbrace{\left\{ \frac{d}{dx_{n-1}} \cdots \frac{d}{dx_1} (n-1)! \prod_{i=1}^{n-1} \left[F(x_i) - F(y_i)\right] \right\}}$$

Proof

$$= \frac{d}{dx_n} n\left[\underset{\downarrow\ f(x_n)}{F(x_n)} - \cancel{F(y_n)}\right] \underbrace{(n-1)!\, f(x_1) \cdots f(x_{n-1})}$$

$$= n(n-1)!\, f(x_n) f(x_1) \cdots f(x_{n-1})$$

$$= n!\, f(x_1) \cdots f(x_n)\, I(x_1 < x_2 < \cdots < x_n)$$

Proof

1.6.4

The distribution of $X_{(i)}$

CDF of $X_{(i)}$

For an i.i.d. sample of continuous random variables $X_1, X_2, \ldots, X_n$ with cdf F, the cdf of $X_{(i)}$ is

$$F_{X_{(i)}}(x) = P(X_{(i)} \le x) = \sum_{j=i}^{n} \binom{n}{j} (1 - F(x))^{n-j} F(x)^j.$$

$$P(X_{(i)} \le x) = P\Big(\{X_{(n)} \le x\} \cup \{X_{(n)} > x, X_{(n-1)} \le x\} \cup \cdots \cup \{X_{(n)} > x, X_{(n-1)} > x, \ldots, X_{(i+1)} > x, X_{(i)} \le x\}\Big)$$

$\leftarrow$ binomial$(n, F(x))$

$$P(X_{(n)} > x, \ldots, X_{(j+1)} > x, X_{(j)} \le x) = \binom{n}{j} F(x)^j (1 - F(x))^{n-j}$$

$$\therefore \quad F_{X_{(i)}}(x) = \sum_{j=i}^{n} \binom{n}{j} F(x)^j (1 - F(x))^{n-j}$$

Proof

Proof (cont)

The PDF of $X_{(i)}$

For an i.i.d. sample of continuous random variables $X_1, X_2, \ldots, X_n$ with cdf F, pdf of f
the pdf of $X_{(i)}$ is

$$f_{X_{(i)}}(x) = \frac{n!}{(i-1)!(n-i)!} F(x)^{i-1} (1 - F(x))^{n-k} f(x).$$

$$f_{X_{(i)}}(x) = \frac{d}{dx} F_{X_{(i)}}(x) \Leftarrow$$

$$= \frac{d}{dx} \sum_{j=i}^{n} \binom{n}{j} F(x)^j (1-F(x))^{n-j}$$

$$= \sum_{j=i}^{n} \frac{d}{dx} \left\{ \binom{n}{j} F(x)^j [1-F(x)]^{n-j} \right.$$

$$= \sum_{j=i}^{n} \binom{n}{j} j [F(x)]^{j-1} f(x) [1-F(x)]^{n-j} \leftarrow$$

$$- \sum_{j=i}^{n} \binom{n}{j} F(x)^j (n-j)(1-F(x))^{n-j-1} f(x)$$

Proof

$$i\binom{n}{i}F(x)^{i-1}f(x)[1-F(x)]^{n-i}$$

$$+\sum_{j=i+1}^{n}\binom{n}{j}j[F(x)]^{j-1}f(x)[1-F(x)]^{n-j}$$

$$-\sum_{j=i}^{n-1}\binom{n}{j}F(x)^{j}(n-j)(1-F(x))^{n-j-1}f(x)$$

Let $k=j-1$

$$j\binom{n}{j} = j\frac{n!}{j!(n-j)!} = \frac{n!}{(j-1)!(n-j)!} = \frac{n!}{k!(n-k-1)!} = (n-k)\frac{n!}{k!(n-k)!} = (n-k)\binom{n}{k}!$$

$$(n-k)\binom{n}{n-k}!$$

$$\sum_{k=i}^{n-1}(n-k)\binom{n}{k}F(x)^{k}(1-F(x))^{n-k-1}f(x)$$

$$-\sum_{j=i}^{n-1}(n-j)\binom{n}{j}F(x)^{j}(1-F(x))^{n-j-1}f(x)$$

Proof (cont)

$$f_{X_{(i)}}(x) = i\binom{n}{i} F(x)^{i-1}[1-F(x)]^{n-i} f(x)$$

$$= i\,\frac{n!}{i!\,(n-i)!} F(x)^{i-1}[1-F(x)]^{n-i} f(x)$$

$$= \frac{n!}{(i-1)!\,(n-i)!} F(x)^{i-1}[1-F(x)]^{n-i} f(x)$$

$$= n\,\frac{(n-1)!}{(i-1)!\,(n-i)!} F(x)^{i-1}[1-F(x)]^{n-i} f(x)$$

$$= n\binom{n-1}{i-1} F(x)^{i-1}[1-F(x)]^{(n-1)-(i-1)}$$

Proof (cont)

1.6.5

The joint distribution of $X_{(i)}$ and $X_{(j)}$

The joint pdf of $X_{(i)}$ and $X_{(j)}$

For an i.i.d. sample of continuous random variables $X_1, X_2, \ldots, X_n$ with cdf F, the joint pdf of $(X_{(i)}, X_{(i)})$ is

$$
\begin{aligned}
& f_{X_{(i)},X_{(j)}}(x_i, x_j) \\
& = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} F(x_i)^{i-1} f(x_i) \\
& \quad \times [F(x_j) - F(x_i)]^{j-i-1} f(x_j) [1 - F(x_j)]^{n-j} I_{(-\infty, x_j)}(x_i).
\end{aligned}
$$