Bayesian Statistics

Basic Distributions

Brani Vidakovic
Professor
School of Industrial and Systems Engineering

Bayesian Statistics

Basic Distributions

Brani Vidakovic

Professor

School of Industrial and Systems Engineering

Numerical Characteristics
of Random Variables

Bayesian Statistics

Basic Distributions

Brani Vidakovic
Professor
School of Industrial and Systems Engineering

Joint and Conditional Distributions

Before We Begin…

In this unit:

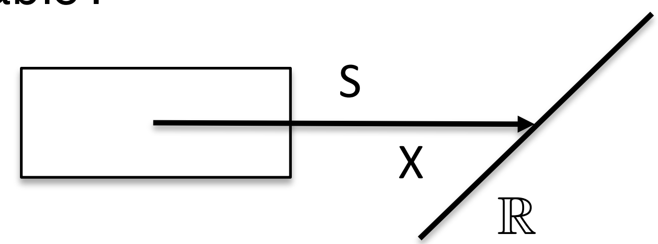
- Basic Distributions
- Bayes's Theorem
- Homework 2
- Conjugate Classes
- Prior Elicitation
- Empirical Bayes
- Homework 3

Basic Distributions

- Models and Parameters
- Numerical Characteristics
- Joint and Conditional Distributions
- Worked Examples

Models and Parameters

- Random variable is the simplest statistical model
- What is random variable?



X: Mapping from sample space S to real numbers

<u>Example</u>: Flip a fair coin. If head, pay a $1, if tail, get a $1. Your gain in one flip is a random variable.

- Random variables: Discrete and continuous (depending on realizations)
- Fully determined by distributions
  - **Discrete**: probability mass function (pmf)

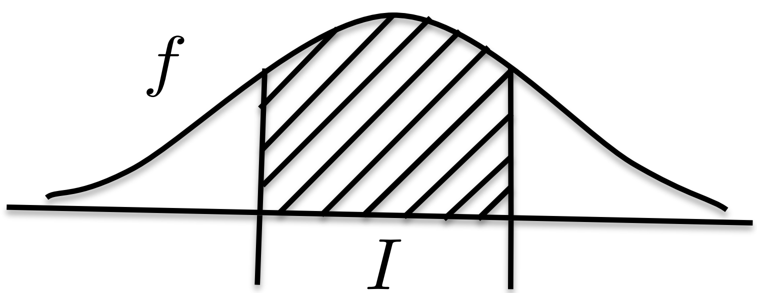
X	x_1	x_2	...	x_n	...
Probs	p_1	p_2	...	p_n	...

$$P(X = x_i) = p_i; \ \sum p_i = 1.$$

Example: Gain in one flip X:

X	-1	1
Prob	$\frac{1}{2}$	$\frac{1}{2}$

  - **Continuous**: Realizations are all points in an interval. Need density: $f(x)$



$$P(X \in I) = \int_I f(x)dx$$

Discrete Distributions

- Bernoulli: $\begin{array}{c|cc} X & 0 & 1 \\ \hline \text{Prob} & q & p \end{array}$, $q = 1 - p$, $0 \leq p \leq 1$

  p — Parameter; $\boxed{X \sim \text{Ber}(p)}$

- Binomial: $\begin{array}{c|cccc} X & 0 & 1 & \ldots & n \\ \hline \text{Prob} & p_0 & p_1 & \ldots & p_n \end{array}$, $p_k = P(X = k) = \binom{n}{k} p^k q^{n-k}, \ q = 1 - p, \ k = 0, 1, \ldots, n$

  $n, \ p$ — Parameters; $\boxed{X \sim \text{Bin}(n, p)}$

- Poisson: $\begin{array}{c|ccccc} X & 0 & 1 & \ldots & n & \ldots \\ \hline \text{Prob} & p_0 & p_1 & \ldots & p_n & \ldots \end{array}$, $p_k = P(X = k) = \dfrac{\lambda^k}{k!} e^{-\lambda}, \ \lambda > 0, \ k = 0, 1, 2, \ldots, n, \ldots$

  λ — Parameter; $\boxed{X \sim \text{Pois}(\lambda)}$

Continuous Distributions

- <u>Uniform:</u> $f(x) = \begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & \text{else} \end{cases}, \quad a,\ b$ — Parameters; $\boxed{X \sim \mathrm{U}(a,b)}$
- <u>Exponential:</u> $f(x) = \lambda e^{-\lambda x},\ \ \lambda > 0,\ \ x \ge 0$, or $f(x) = \dfrac{1}{\mu} e^{-x/\mu},\ \ \mu > 0,\ \ x \ge 0$

  λ — Parameter (rate); $\boxed{X \sim \mathrm{Exp}(\lambda)}$, or μ — Parameter (scale); $\boxed{X \sim \mathrm{Exp}(\mu)}$
- <u>Beta:</u> $f(x) = \dfrac{1}{B(a,b)} x^{a-1}(1-x)^{b-1},\ \ 0 \le x \le 1$

  B: Beta function, defined via integral $\displaystyle\int_0^1 x^{a-1}(1-x)^{b-1}dx$;
  $a,\ b$ — Parameters; $\boxed{X \sim \mathrm{Be}(a,b)}$

Normal Distribution

- <u>Normal:</u> $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \ -\infty < x < +\infty, \ -\infty < \mu < +\infty, \ \sigma > 0$

  $\mu, \ \sigma^2$ — Parameters; $\boxed{X \sim N(\mu, \sigma^2)}$

- Let $X \sim f(x)$
  - $f(x-\mu) \leftarrow \mu$ is location shift parameter
  - $\frac{1}{\sigma} f(\frac{x}{\sigma}) \leftarrow \sigma$ is scale parameter
- If $f(x)$ depends on parameter θ, we write $f(x|\theta)$.

Cumulative Distribution Function (cdf)

pmf: Probability mass function (discrete)
pdf: Probability density function (continuous)

cdf: Probability Distribution (cumulative)

$$F(x) \stackrel{\text{def}}{=} P(X \leq x) = \begin{cases} \sum_{x_i \leq x} P(X = x_i) \\ \int_{-\infty}^{x} f(x)dx \end{cases}$$

<u>Example</u>: $X \sim \mathrm{Exp}(\lambda)$, $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$;

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{0}^{x} \lambda e^{-\lambda t} dt = -e^{-\lambda t}|_{0}^{x} = -e^{-\lambda x} - (-1) = 1 - e^{-\lambda x}, x \geq 0$$

Numerical Characteristics of Random Variables

Numerical Characteristics

- **Expectation, moments:**

Expectation

$$EX = x_1p_1 + x_2p_2 + ... + x_np_n + ... \text{ (for discrete)}$$
$$= \int_{\mathbb{R}} x \cdot f(x)dx \text{ (for continuous)}$$

k^{th} moment

$$EX^k = x_1^kp_1 + x_2^kp_2 + ... + x_n^kp_n + ... \text{ (for discrete)}$$
$$= \int_{\mathbb{R}} x^k \cdot f(x)dx \text{ (for continuous)}$$

General function φ

$$E\varphi(X) = \varphi(x_1)p_1 + \varphi(x_2)p_2 + ... + \varphi(x_n)p_n + ... \text{ (for discrete)}$$
$$= \int_{\mathbb{R}} \varphi(x) \cdot f(x)dx \text{ (for continuous)}$$

Variance and Standard Deviation

- **Variance:** $\mathrm{Var}(X) \stackrel{\text{def}}{=} E(X - EX)^2 = EX^2 - (EX)^2$
- **Standard deviation:** $\sigma_x = \sqrt{\mathrm{Var}(X)}$

Example: If $X \sim \mathrm{Exp}(\lambda)$, $\mathrm{Var}(X) = \dfrac{1}{\lambda^2};\ \sigma_x = \dfrac{1}{\lambda}.$

Quantiles

- ξ_p is p^{th} quantile of distribution F if $F(\xi_p) = p$.
    - $\xi_p = \inf\{x | \sum_{x_i \le x} p_i \ge p\}$ (discrete)
    - $\int_0^{\xi_p} f(x)dx = p$ or $F^{-1}(p) = \xi_p$ (continuous)

<u>Example:</u> If $X \sim \mathrm{Exp}(\lambda)$; $F(x) = 1 - e^{-\lambda x},\ \lambda > 0,\ x \ge 0$

$$F(\xi_p) = p,\ \text{and}\ \xi_p = -\frac{1}{\lambda}\log(1-p),\ 0 \le p \le 1.$$

- Median $\stackrel{\text{def}}{=} \xi_{1/2}$ (50% percentile, 0.5-quantile)
- $Q_1,\ Q_3 \stackrel{\text{def}}{=} \xi_{1/4},\ \xi_{3/4}$ (first and third quantiles)
- Mode: Most frequent/likely value
    - Continuous: Value that maximizes density
    - Discrete: Value x_i for which $p_i = P(X = x_i)$ is maximum.

<u>Example:</u> $X \sim \mathrm{Exp}(\lambda)$, median = $\dfrac{\log 2}{\lambda}$, mode = 0 .

Joint and Conditional Distributions

Vector r.v. $X = (X_1, X_2, ..., X_n)$

- **Joint distribution**

  pdf, cdf $f(x_1, x_2, ..., x_n), \; F(x_1, x_2, ..., x_n)$

  $X = (X_1, X_2) \leftarrow$ two dimensional case

- **Conditional distribution** $X_1, \;$ given X_2

$$f(x_1|x_2) \stackrel{\text{def}}{=} \frac{f(x_1, x_2)}{f(x_2)};$$

  here $f(x_2) \leftarrow$ marginal for X_2, defined as $\int_{-\infty}^{+\infty} f(x_1, x_2) dx_1$.

Worked Examples

Example 1: Discrete 2-D

X \ Y	-1	0	1
0	0.1	0.3	0.1
1	0.1	0.05	0.1
2	0.05	0.1	0.1

X	Prob
0	0.5
1	0.25
2	0.25

Y	-1	0	1
Prob	0.25	0.45	0.30

X|Y=0	Prob
0	$\frac{0.3}{0.45} = \frac{2}{3}$
1	$\frac{0.05}{0.45} = \frac{1}{9}$
2	$\frac{0.1}{0.45} = \frac{2}{9}$

Y|X=2	-1	0	1
Prob	$\frac{0.05}{0.25} = 0.2$	$\frac{0.1}{0.25} = 0.4$	$\frac{0.1}{0.25} = 0.4$

Example 2: Continuous 2-D

- If $f(x,y) = \frac{1}{\pi} e^{-\frac{1}{2}(x^2 - 2xy + 5y^2)},\ x,\ y \in \mathbb{R}^2$, find marginal distributions for $X,\ Y$ and find Conditional distributions for $X|Y=y$ and $Y|X=x$.

$$
\begin{aligned}
f(x,y) &= \frac{1}{\pi} e^{-\frac{1}{2}(x^2 - 2xy + 5y^2)} \\
&= \frac{1}{\pi} e^{-2y^2} e^{-\frac{1}{2}(x-y)^2} \\
&= \frac{1}{\pi} \cdot e^{-2y^2} \cdot \sqrt{2\pi} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-y)^2}
\end{aligned}
$$

$$
\begin{aligned}
f(y) &= \int_{\mathbb{R}} f(x,y)dx = \sqrt{\frac{2}{\pi}} e^{-2y^2} \cdot \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-y)^2} dx \\
&= \frac{1}{\sqrt{2\pi \cdot \frac{1}{4}}} e^{-\frac{y^2}{2 \cdot \frac{1}{4}}} \cdot 1 \sim N\left(0, \left(\frac{1}{2}\right)^2\right)
\end{aligned}
$$

- $f(x) = \int f(x,y)dy$

$$f(x,y) = \frac{1}{\pi} e^{-\frac{5}{2}\left(y^2 - \frac{2x}{5}y + \frac{x^2}{25} - \frac{x^2}{25} + \frac{x^2}{5}\right)}$$

$$= \frac{1}{\pi} e^{-\frac{5}{2}\cdot\frac{4}{25}x^2} \cdot e^{-\frac{5}{2}(y-\frac{x}{5})^2}$$

$$= \frac{1}{\pi} \cdot e^{-\frac{2}{5}x^2} \cdot \sqrt{2\pi \cdot \frac{1}{5}} \cdot \frac{1}{\sqrt{2\pi \cdot \frac{1}{5}}} \cdot e^{-\frac{1}{2\cdot\frac{1}{5}}\left(y-\frac{x}{5}\right)^2}$$

$$f(x) = \frac{1}{\sqrt{2\pi \cdot \frac{5}{4}}} e^{-\frac{x^2}{e^{2\cdot\frac{5}{4}}}} \cdot \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\cdot\frac{1}{5}}} e^{-\frac{1}{2\cdot\frac{1}{5}}\left(y-\frac{x}{5}\right)^2} \quad \to X \sim N\left(0, \left(\frac{\sqrt{5}}{2}\right)^2\right)$$

Thus, marginals are normal $X \sim N\left(0, \left(\frac{\sqrt{5}}{2}\right)^2\right)$, and $Y \sim N\left(0, \left(\frac{1}{2}\right)^2\right)$.

- Conditionals: By definition $f(x|y) \stackrel{\text{def}}{=} \dfrac{f(x,y)}{f(y)}$

Recall: $P(A\ B) \stackrel{\text{def}}{=} P(A|B)P(B)$
$f(x,y) \stackrel{\text{def}}{=} f(x|y) \cdot f(y)$

$$f(x|y) = \frac{\frac{1}{\pi} e^{-\frac{1}{2}(x^2 - 2xy + 5y^2)}}{\frac{1}{\sqrt{2\pi \cdot \frac{1}{4}}} \cdot e^{-\frac{1}{2} \cdot 4y^2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-y)^2}, \ x \in \mathbb{R}$$

$$\to X|Y = y \sim N(y, 1)$$

$$f(y|x) = \frac{\frac{1}{\pi} e^{-\frac{1}{2}(x^2 - 2xy + 5y^2)}}{\frac{1}{\sqrt{2\pi \cdot \frac{5}{4}}} e^{-\frac{1}{2} \cdot x^2 \cdot \frac{4}{5}}} = \frac{1}{\sqrt{2\pi \cdot \frac{1}{5}}} e^{-\frac{1}{2 \cdot \frac{1}{5}}(y - \frac{x}{5})^2}, \ y \in \mathbb{R}$$

$$\to Y|X = x \sim N\left(\frac{x}{5}, \left(\frac{1}{\sqrt{5}}\right)^2\right)$$

Example 3: Independent 2-D

- If $f(x,y) = 2xe^{-x-2y}, x \geq 0, y \geq 0$, find conditional distributions for $f(x|y)$ and $f(y|x)$.

x and y are separates variables: $f(x,y) = xe^{-x} \cdot 2e^{-2y} = f(x) \cdot f(y)$
$\rightarrow (X,Y), X, Y$ are independent components; because of independence:

$$f(x|y) = f(x),\ f(y|x) = f(y)$$

Summary