# Bayesian Analysis and Poisson Statistics

Stewart Nash

September 29, 2019

## 1  Bayesian Statistics

### 1.1  Conditional Probability

The conditional probability of an event $A$ assuming event $M$ is

$$P(A|M) = \frac{P(A \cap M)}{P(M)}. \tag{1}$$

### 1.2  Total Probability Theorem

If $\mathbf{U} = [A_1, \ldots, A_n]$ is a partition of $S$ and $B$ is an arbitrary event, then

$$P(B) = P(B|A_1)P(A_1) + \cdots + P(B|A_n)P(A_n). \tag{2}$$

### 1.3  Bayes' Theorem

The *a priori* probability is $P(A_i)$ and the *a posteriori* probability is $P(A_i|B)$. Bayes' theorem is as follows

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B|A_1)P(A_1) + \cdots |P(B|A_n)P(A_n)}. \tag{3}$$

### 1.4  Independence

Two events $A$ and $B$ are independent if

$$P(A \cap B) = P(A)P(B). \tag{4}$$

Independence can be defined inductively for $n > 2$. The independence of multiple events can be written as follows

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i) \tag{5}$$

$$P(A_i \cap A_j) = P(A_i)P(A_j) \quad i, j \in I, \quad i \neq j \tag{6}$$

**Problem 1.1** *(2-8) If $A \subset B$, $P(A) = 1/4$, and $P(B) = 1/3$, find $P(A|B)$ and $P(B|A)$.*

**Solution 1.1**

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)} = \frac{1/4}{1/3} = 3/4 = 0.75$$

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1$$

**Problem 1.2** *(2-9) Show that $P(AB|C) = P(A|BC)P(B|C)$ and $P(ABC) = P(A|BC)P(B|C)P(C)$.*

**Solution 1.2** *(a)*

$$P(ABC) = P(AB|C)P(C)$$
$$P(A|BC)P(BC) = P(AB|C)P(C)$$
$$P(A|BC)P(B|C)P(C) = P(AB|C)P(C)$$
$$P(AB|C) = P(A|BC)P(B|C)$$

*(b)*

$$P(ABC) = P(A|BC)P(BC)$$
$$= P(A|BC)P(B|C)P(C)$$

**Problem 1.3** *(2-12) A call occurs at time $t$, where $t$ is a random point in the interval $(0, 10)$.*
*(a) Find $P\{6 \leq t \leq 8\}$.*
*(b) Find $P\{6 \leq t \leq 8 | t > 5\}$.*

**Solution 1.3** *(a) $P\{6 \leq t \leq 8\} = \frac{8-6}{10} = \frac{1}{5} = 0.2$*
*(b)*

$$P\{t > 5\} = 1/2 = 0.5$$
$$P\{6 \leq t \leq 8 | t > 5\} = \frac{P\{6 \leq t \leq 8, t > 5\}}{P\{t > 5\}}$$
$$= \frac{P\{6 \leq t \leq 8\}}{P\{t > 5\}}$$
$$= \frac{1/5}{1/2} = 2/5 = 0.4$$

**Problem 1.4** *(2-14) The events $A$ and $B$ are mutually exclusive. Can they be independent?*

**Solution 1.4**
$$P(A \cap B) = P(\emptyset) = 0 \neq P(A)P(B)$$

*No, they cannot be independent.*

**Problem 1.5** *(2-24) Box 1 contains 1000 bulbs of which 10% are defective. Box 2 contains 2000 bulbs of which 5% are defective. Two bulbs are picked from a randomly selected box. (a) Find the probability that both bulbs are defective. (b) Assuming that both are defective, find the probability that they come from box 1.*

**Solution 1.5** *Denote the probability of a defective bulb on the first and second draw as $P(D) = P(D_1 D_2)$ and denote the probability of drawing from box $i$ as $P(B_i)$. From the total probability theorem, we obtain*

$$P(D) = \frac{P(D|B_i)P(B_i)}{P(B_i|D)} \, .$$

*Bayes' theorem yields*

$$P(B_i|D) = \frac{P(D|B_i)P(B_i)}{\sum_k P(D|B_k)P(B_k)} \, .$$

*Combining these, we get*

$$P(D) = \sum_k P(D|B_k)P(B_k)$$

$$P(D) = P(D|B_1)P(B_1) + P(D|B_2)P(B_2)$$

*(a) Therefore, we have*

$$P(D|B_1) = P(D_1 D_2|B_1)$$
$$= (0.1)(99/1000) = 0.0099$$
$$P(D|B_2) = P(D_1 D_2|B_2)$$
$$= (0.05)(99/2000) = 0.002475$$
$$P(B_1) = P(B_2) = 1/2 = 0.5$$
$$P(D) = (0.0099)(0.5) + (0.002475)(0.5) = 0.0061875 \, .$$

*(b)*

$$P(B_1|D) = \frac{P(D|B_1)P(B_1)}{\sum_k P(D|B_k)P(B_k)}$$
$$= \frac{P(D|B_1)P(B_1)}{P(D)}$$
$$= (0.0099)(0.5)/0.0061875$$
$$= 0.8$$

# 2 Random Variables

## 2.1 Conditional Probability

The conditional distribution $F(x|M)$ of a random variable $\mathbf{x}$, assuming $M$ is defined as the conditional probability of the event $\{\mathbf{x} \leq x\}$:

$$F(x|M) = P\{\mathbf{x} \leq x|M\} = \frac{P\{\mathbf{x} \leq x, M\}}{P(M)} . \tag{7}$$

The conditional density $f(x|M)$ is the derivative of $F(x|M)$:

$$f(x|M) = \frac{dF(x|M)}{dx} = \lim_{\Delta x \to 0} \frac{P\{x \leq \mathbf{x} \leq x + \Delta x|M\}}{\Delta x} . \tag{8}$$

## 2.2 Total Probability Theorem

$$\int_{-\infty}^{\infty} P(A|\mathbf{x} = x) f(x)\, dx = P(A) \tag{9}$$

## 2.3 Bayes' Theorem

$$f(x|A) = \frac{P(A|\mathbf{x} = x)}{P(A)} f(x) = \frac{P(A|\mathbf{x} = x) f(x)}{\int_{-\infty}^{\infty} P(A|\mathbf{x} = x) f(x)\, dx} \tag{10}$$

**Problem 2.1** *(4-10)If $\mathbf{x}$ is $N(0,2)$ find (a) $P\{1 \leq \mathbf{x} \leq 2\}$ and (b) $P\{1 \leq \mathbf{x} \leq 2|\mathbf{x} \geq 1\}$.*

**Solution 2.1** *(a) We have $\mathbf{x} \sim N(\mu, \sigma^2)$ with $\mu = 0$ and $\sigma = \sqrt{2}$.*

$$P\{a \leq \mathbf{x} \leq b\} = G\left(\frac{b - \mu}{\sigma}\right) - G\left(\frac{a - \mu}{\sigma}\right)$$
$$P\{1 \leq \mathbf{x} \leq 2\} = G(\sqrt{2}) - G(\sqrt{2}/2)$$

*(b) We have $\mathbf{x} \sim N(\mu, \sigma^2)$ with $\mu = 0$ and $\sigma = \sqrt{2}$.*

$$P\{1 \leq \mathbf{x} \leq 2|\mathbf{x} \geq 1\} = \frac{P\{\mathbf{x} \geq 1|1 \leq \mathbf{x} \leq 2\}P\{1 \leq \mathbf{x} \leq 2\}}{P\{\mathbf{x} \geq 1\}}$$
$$= \frac{P\{1 \leq \mathbf{x} \leq 2\}}{P\{\mathbf{x} \geq 1\}}$$
$$= \frac{G(\sqrt{2}) - G(\sqrt{2}/2)}{1 - G(\sqrt{2}/2)}$$

**Problem 2.2** *(4-21)The probability of heads of a random coin is a random variable $\mathbf{p}$ uniform in the interval $(0,1)$. (a) Find $P\{0.3 \leq \mathbf{p} \leq 0.7\}$. (b) The coin is tossed 10 times and heads shows 6 times. Find the a posteriori probability that $\mathbf{p}$ is between 0.3 and 0.7.*

4

**Solution 2.2** *(a)* $P\{0.3 \leq \mathbf{p} \leq 0.7\} = 0.7 - 0.3 = 0.4$
*(b) If A is k heads in n specific tosses*

$$f_p(p|A) = \frac{(n+1)!}{(n-k)!k!}p^k q^{n-k}$$

$$F_p(b|A) - F_p(a|A) = \int_a^b f_p(p|A)\,dp$$

$$= \frac{(n+1)!}{(n+1-k)!(k+1)!}p^{k+1}q^{n+1-k}|_{p=a}^b$$

*where $b = 0.7$, $a = 0.3$, $n = 10$ and $k = 6$.*

# 3  Bayesian Network

A *Bayesian network* is a *belief network* or a *probabilistic network*. A *node* is a random variable and an *arc* is an influence. The Bayesian network is a graphical model that consists of nodes and arcs. It is a *directed acyclic graph*. The directed acylic graph may be turned into a *junction tree*. The *belief propagation* algorithm can make inferences from a tree. A link (arc?) does not imply causality (between nodes), only a direct influence.

# 4  Bayesian Estimation

Given an iid (independent and identically distributed) sample $X = \{x_i\}_{i=1}^N$, assume that $x_i t$ are instances drawn from some probability density family $p(\mathbf{x}; \theta)$ defined up to parameter $\theta$:

$$x_i \sim p(\mathbf{x}; \theta) \tag{11}$$

We would like to find $\theta$ that make sampling $x_i$ from $p(\mathbf{x}; \theta)$ as likely as possible. Because the $x^t$ are independent, the likelihood of sample $X$ given the parameter $\theta$ is the product of the likelihoods of the individual points

$$l(\theta) \equiv p(X|\theta) = \prod_{i=1}^n p(x_i|\theta) \tag{12}$$

When we have prior information on the possible value range of a parameter $\theta$, we model $\theta$ as a random variable and define a prior density $p(\theta)$. We combine this with the sample data – likelihood density $p(X|\theta)$ – using Bayes' rule to give us the posterior density of theta

$$p(\theta|X) = \frac{p(X|\theta)p(\theta)}{p(X)} = \frac{p(X|\theta)p(\theta)}{\int p(X|\theta')p(\theta')\,d\theta'} \tag{13}$$

To estimate the density at $x$ we have

$$p(x|X) = \int p(x, \theta|X) \, d\theta$$

$$= \int p(x|\theta)p(\theta|X) \, d\theta$$

The Bayes' estimator is the expected value of the posterior density

$$E[\theta|X] = \int \theta p(\theta|X) \, d\theta \qquad (14)$$

# 5   Bernoulli Trials

The number of *permutations* of $n$ objects is $n!$. The number of *permutations* of $n$ objects taken $k$ at a time is

$$\frac{n!}{(n-k)!} \qquad (15)$$

The total *combinations* of $n$ objects taken $k$ at a time is

$$\binom{n}{k} = \frac{n!}{(n-k)!k!} \qquad (16)$$

We are given an experiment $S$ and an event $A \in S$ with

$$P(A) = p$$
$$P(\bar{A}) = q$$
$$q + p = 1$$

Repeat the experiment $n$ times to obtain (the product space) $S^n$. $p_n(k)$ is the probability that $A$ occurs $k$ times

$$p_n(k) = \binom{n}{k} p^k q^{n-k} \qquad (17)$$

**Problem 5.1** *(3-2) A pair of dice is rolled 50 times. Find the probability of obtaining double six at least three times.*

**Solution 5.1**

$$p = 1/6$$
$$q = 5/6$$
$$p_{50}(6) = \binom{50}{6} p^6 q^{44} \approx 0.111753$$

**Problem 5.2** *(3-3) A pair of fair dice is rolled 10 times. Find the probability that "seven" will show at least once.*

**Solution 5.2**

$$p = 1 - \left(\frac{5}{6}\right)^{10} \approx 0.838494$$

**Problem 5.3** *(3-4) A coin with $p\{h\} = p = 1 - q$ is tossed $n$ times. Show that the probability that the number of heads is even equals $0.5[1 + (q - p)^n]$.*

**Solution 5.3**

$$\sum_{k=1}^{n/2} p_n(2k) = \sum_{k=1}^{n/2} \binom{n}{2k} p^{2k} q^{n-2k}$$

$$0.5[1 + (q - p)^n] = \frac{1}{2}\left(1 + \sum_{k=0}^{n} \binom{n}{k}(-p)^k q^{n-k}\right)$$

$$= \frac{1}{2}\left(\sum_{k=0}^{n}\binom{n}{k}p^k q^{n-k} + \sum_{k=0}^{n}(-1)^k\binom{n}{k}p^k q^{n-k}\right)$$

$$= \frac{1}{2}\sum_{k=0}^{n}(1 + (-1)^k)\binom{n}{k}p^k q^{n-k}$$

$$= \frac{1}{2}\sum_{k=0}^{\lfloor n/2\rfloor} 2\binom{n}{2k}p^{2k}q^{n-2k}$$

$$= \sum_{k=0}^{\lfloor n/2\rfloor} p_n(2k)$$

**Problem 5.4** *Three dice are rolled and the player may bet on any one of the face values 1, 2, 3, 4, 5 and 6. If the players number appears on one, two, or all three dice, the player receives respectively one, two or three times his original stake plus his own money back. Determine the expected loss per unit stake for the player.*

**Solution 5.4**

$$\frac{74}{216}$$

# 6   Poisson Process

A discrete random variable is a Bernoulli random variable if it has a binary outcome. $\mathbf{x}$ is Bernoulli distributed if

$$P\{\mathbf{x} = 1\} = p \tag{18}$$

$$P\{\mathbf{x} = 0\} = q = 1 - p \tag{19}$$

A binomial random variable $\mathbf{y}$ represents the number of favorable outcomes in an independent trial of $n$ Bernoulli experiments

$$P\{\mathbf{y} = k\} = \binom{n}{k}p^k q^{(n-k)} \quad k \in \mathbb{N}, \, 0 \le k \le n \tag{20}$$

A Poisson distribution represents the number of occurrences of a rare event in a large number of trials. If $\mathbf{x}$ is a Poisson random variable with parameter $\lambda$

$$P\{\mathbf{x} = k\} = e^{-\lambda}\frac{\lambda^k}{k!} = p_k \quad k \in \mathbb{N} \tag{21}$$

It follows that

$$\frac{p_{k-1}}{p_k} = \frac{k}{\lambda} \tag{22}$$

# 7 Poisson Theorem

If $n \to \infty$ and $p \to 0$ such that $np \to \lambda$ then

$$\frac{n!}{k!(n-k)!}p^k q^{n-k}\overrightarrow{n \to \infty}e^{-\lambda}\frac{\lambda^k}{k!} \quad , k \in \mathbb{N} \tag{23}$$

$$P\{a \le k \le b\} \simeq e^{-np}\sum_{k=a}^{b}\frac{(np)^k}{k!} \tag{24}$$

## 7.1 Generalization of Poisson Theorem

Suppose that $A_1, \ldots, A_{m+1}$ are the $m+1$ events of a partition with $P\{A_i\} = p_i$. If $np_i \to a_i$ for $i \le m$, then

$$\frac{n!}{k_1! \cdots k_{m+1}!}p_1^{k_1} \cdots p_{m+1}^{k_{m+1}}\overrightarrow{n \to \infty}\frac{e^{-a_1}a_1^{k_1}}{k_1!} \cdots \frac{e^{-a_m}a_n^{k_m}}{k_m!} \tag{25}$$

# 8 Random Poisson Points

If we randomly place points on the $t$-axis and take two arbitrary intervals, $(t_1, t_2)$ and $(t_3, t_4)$, of length $t_2 - t_1 = t_a$ and $t_4 - t_3 = t_b$, then this experiment generates random Poisson points if

1.
$$P\{k_a \text{ in } t_a\} = e^{-\lambda t_a}\frac{(\lambda t_a)^{k_a}}{k_a!} \tag{26}$$

   for $(t_2, t_1)$

2. and for the non-overlapping intervals $(t_1, t_2)$ and $(t_3, t_4)$ the events $\{k_a \text{ in } (t_1, t_2)\}$ and $\{k_b \text{ in } (t_3, t_4)\}$ are independent.

## 8.1 Density of Poisson Points

The parameter $\lambda$ can be interpreted as the density of Poisson points. For an apparently/optionally non-uniform density $\lambda(t) \ge 0$ we have

$$P\{k \text{ in } (t_1, t_2)\} = \exp\left(-\int_{t_1}^{t_2}\lambda(t)\,dt\right)\frac{\left(\int_{t_1}^{t_2}\lambda(t)\,dt\right)^k}{k!} \tag{27}$$