

Probability and Random Processes

Irvin Avalos

1 Information Theory

Defintion 1.1 (Entropy).

$$H(X) = \mathbb{E} \left[-\log_2 p(x) \right] = \sum_{x \in X} -p(x) \log_2 p(x)$$

If $X \sim B(p)$, then

$$H(X) = p - \log_2 p + (1 - p) \log_2 (1 - p) \triangleq h(p)$$

called the **binary entropy function**.

Defintion 1.2 (Joint Entropy).

$$H(X, Y) = \mathbb{E} \left[-\log_2 p(x, y) \right] = \sum_x \sum_y -p_{x,y}(x, y) \log_2 p(x, y)$$

If X, Y are independent, then $H(X, Y) = H(X) + H(Y)$.

Proof.

$$\begin{aligned} H(X, Y) &= \mathbb{E} \left[-\log_2 p(x, y) \right] \\ &= \mathbb{E} \left[-\log_2 p(x) p(y) \right] \\ &= \mathbb{E} \left[-\log_2 p(x) - \log_2 p(y) \right] \\ &= \mathbb{E} \left[-\log_2 p(x) \right] + \mathbb{E} \left[-\log_2 p(y) \right] \\ &= H(X) + H(Y) \end{aligned}$$

□

Defintion 1.3 (Conditional Entropy).

$$H(Y | X) = \mathbb{E}_{XY} \left[-\log_2 p(y | x) \right] H(X, Y) - H(X)$$

Defintion 1.4 (Mutual Information).

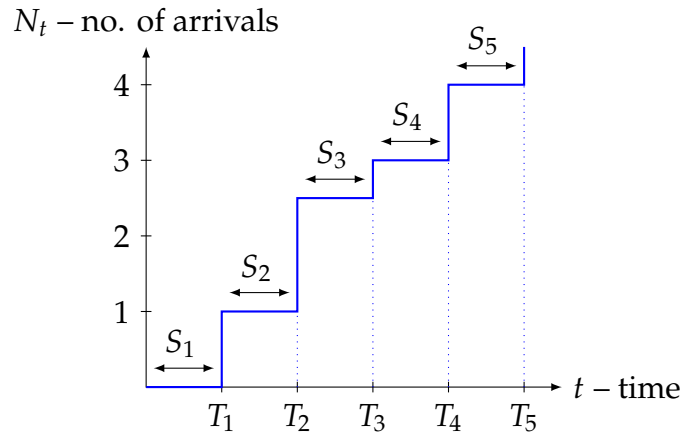
$$\begin{aligned} I(X; Y) &\triangleq H(X) - H(X | Y) \\ &= H(Y) - H(Y | X) \\ &= H(X) + H(Y) - H(X, Y) \end{aligned}$$

Theorem 1.5 (Asymptotic Equipartition). If X_1, X_2, \dots, X_n are iid $\sim p(X)$, then

$$\frac{-\log_2 p(X_1, X_2, \dots, X_n)}{n} \xrightarrow{p} H(X).$$

2 Poisson Process

A Poisson process is the continuous time analog of “coin flipping” (or Bernoulli) processes. This makes it a good model for arrival processes: photons hitting a detector, packets in a network, number of emails per hour, etc.



Each T_i for $i = 1, 2, 3, 4, 5$ represents an arrival and generally, each arrival time is defined as

$$T_n = \sum_{i=1}^n S_i$$

where the interarrival times $S_1, S_2, \dots, S_n \stackrel{\text{iid}}{\sim} \text{Exponential}(\lambda)$. Thus, every S_i has the probability density function

$$f_{S_i}(t) = \lambda e^{-\lambda t}; \quad t > 0; \quad i = 1, 2, 3, \dots$$

and cumulative distribution function

$$F_{S_i}(t) = 1 - e^{-\lambda t}; \quad i = 1, 2, 3, \dots$$

Defintion 2.1 (Number of Arrivals).

$$N_t = \begin{cases} \max_{n \geq 1} \{n \mid T_n \leq t\} & t \geq 0 \\ 0 & t < T_1 \end{cases}$$

- Recall: $\text{Exponential}(\lambda)$ is a memoryless RV

1. $F_\tau(t) = \begin{cases} 1 - e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases} \Rightarrow f_\tau(t) = \lambda e^{-\lambda t}$
2. $E[\tau] = \frac{1}{\lambda}$ and $\text{Var}(\tau) = \frac{1}{\lambda^2}$
3. $P(\tau > t + s \mid \tau > s) = P(\tau > t)$
4. $P(\tau \leq t + \epsilon \mid \tau > t) = \lambda\epsilon + o(\epsilon); \lim_{\epsilon \rightarrow 0} \frac{o(\epsilon)}{\epsilon} = 0$

Proof.

$$\begin{aligned}
P(\tau > t + \epsilon \mid \tau > t) &= P(\tau > \epsilon) \\
&= e^{-\lambda\epsilon} \\
&= 1 - \lambda\epsilon + o(\epsilon) \\
\therefore P(\tau \leq t + \epsilon \mid \tau > t) &= 1 - P(\tau > t + \epsilon \mid \tau > t) \\
&= \lambda\epsilon + o(\epsilon)
\end{aligned}$$

□

- $P(1 \text{ arrival}) = \lambda\epsilon + o(\epsilon)$
- $P(2 \text{ arrivals}) = 1 - \lambda\epsilon + o(\epsilon)$
- $P(3 \text{ arrivals}) = o(\epsilon)$

We can only have one arrival in a particular subinterval with high probability $\lambda\epsilon + o(\epsilon)$. On the other hand, there are zero arrivals with probability $1 - \lambda\epsilon + o(\epsilon)$.

This means that a Poisson process has independent and stationary increments, i.e., for any $0 \leq t_1 \leq t_2 \leq \dots \leq t_n \leq \dots$; if you look at the number of arrivals in t_{n+1} and t_n , $\{N_{t_{n+1}} - N_{t_n}\}$ are independent and the distribution depends only on the size of the interval $(t_{n+1} - t_n)$.

Theorem 2.2. Let $N_t := \text{number of arrivals in } (0, t)$, then

$$P(N_t = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}; \quad k = 0, 1, 2, \dots$$

Proof.

$$P(N_t = k) = \int_{t_1} \int_{t_2} \dots \int_{t_k} f_{T_1, T_2, \dots, T_k \mid N_t = k}(t_1, t_2, \dots, t_k \mid N_t = k) dt_1 dt_2 \dots dt_k$$

We constrain the region of integration by letting $S = 0 \leq t_1 \leq t_2 \leq \dots \leq t_k \leq t \leq t_{k+1}$. Shifting focus to the k^{th} dimensional pdf,

$$f_{T_1, \dots, T_k \mid N_t = k}(t_1, \dots, t_k \mid N_t = k) = P(T_1 \in (t_1, t_1 + dt_1), \dots, T_k \in (t_k, t_k + dt_k), T_{k+1} > t)$$

where we leverage the fact that $P(T_{k+1} \in (t_{k+1}, t)) = P(T_{k+1} > t)$

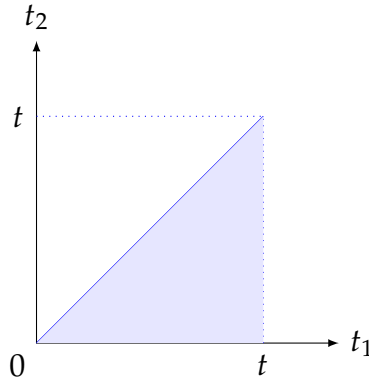
$$\begin{aligned}
&\Rightarrow P(T_1 \in (t_1, t_1 + dt_1)) \times \dots \times P(T_k \in (t_k, t_k + dt_k)) \times P(T_{k+1} > t) \\
&= P(S_1 \in (t_1, t_1 + dt_1)) \times \dots \times P(S_k \in (t_k - t_{k-1}, t_k - t_{k-1} + dt_k)) \times P(S_{k+1} > t - t_k) \\
&= (\lambda e^{-\lambda t_1} dt_1) \dots (\lambda e^{-\lambda(t_k - t_{k-1})} dt_k) e^{-\lambda(t - t_k)} \\
&= \lambda^k e^{-\lambda t}
\end{aligned}$$

The result $\lambda^k e^{-\lambda t}$ is a constant to the k^{th} dimensional integral, so barring any constraints on t_1, t_2, \dots, t_k , $\text{Volume}(S) = t^k$. But, we must respect the fact that $0 \leq t_1 \leq t_2 \leq \dots \leq t_k$. By symmetry, every permutation of (t_1, t_2, \dots, t_k) has equal volume

$$\text{Volume}(S) = \frac{t^k}{k!} \Rightarrow P(N_t = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}; \quad k = 0, 1, 2, \dots$$

□

- Conditioned on the number of arrivals being k in an interval, the density is uniform
- Intuition behind symmetry argument: $k = 2$



From the image above we know that $P(t_1 \leq t_2) + P(t_2 < t_1) = 1$. Thus, $P(t_1 < t_2) = \frac{1}{2} = \frac{1}{2!}$ which generalizes to k dimensions.

Defintion 2.3 (Poisson Merging). Merging two or more independent Poisson processes yields $PP(\lambda_1 + \lambda_2 + \dots + \lambda_n)$

Defintion 2.4 (Poisson Splitting). Let $N \sim PP(\lambda)$ such that N is split into N_1 and N_2 . For each arrival we flip a coin independently with probability $H = p$. If the coin lands on heads, send the arrival to the N_1 queue otherwise it is sent to the N_2 queue.