EMP 214 - Probability cheat sheet[1]

1 Prerequisites

$$
\begin{aligned}
&1 + a + a^2 + a^3 + \cdots = \frac{1}{1-a} \\
&a + a^2 + a^3 + a^4 + \cdots = \\
&a(1 + a + a^2 + a^3 + \cdots) = \frac{a}{1-a} \\
&1 + 2a + 3a^2 + 4a^3 + \cdots = \frac{1}{(1-a)^2}
\end{aligned}
$$

2 Axioms of Probability

- Non negativity : ensures that probability is never negative.
  $\mathbb{P}[A] \geq 0$
- Normalization : ensures that probability is never greater than 1.
  $\mathbb{P}[\Omega] = 1$
- Additivity : allows us to add probabilities when two events do not overlap.

$$\mathbb{P}\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} \mathbb{P}[A_i]$$

2.1 Unions of Two Non-Disjoint Sets

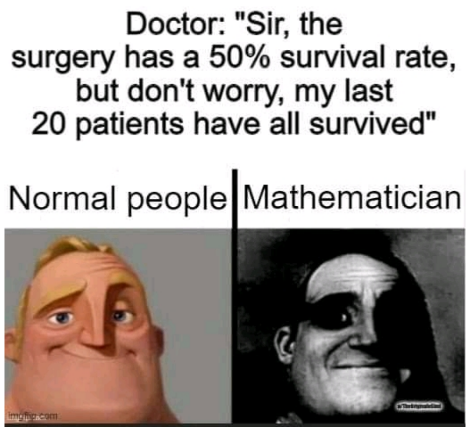
$$\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B]$$

3 Contional probability

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$$

- Independence : Two events A and B are independent if :

$$
\begin{aligned}
&\mathbb{P}[A|B] = \mathbb{P}[A] \\
\text{or} \qquad &\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]
\end{aligned}
$$



[1] Taha Ahmed

- Note:

$$\text{Disjoint} \not\Leftrightarrow \text{Independent}$$

If A and B are disjoint, then $A \cap B = \phi$. This only implies that $\mathbb{P}[A \cap B] = 0$.

- Law of Total Probability :

$$\mathbb{P}[A] = \sum_{i=1}^{n} \mathbb{P}[A|B_i]\mathbb{P}[B_i]$$

- Bayes' rule

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[B|A]\mathbb{P}[A]}{\mathbb{P}[B]}$$

4 Series and Parallel Circuits

- Series devices:

$$
\begin{aligned}
\mathbb{P}[\text{Circuit } \textbf{operates}] = \mathbb{P}&[\text{device 1 } \textbf{operates}] \\
&\times \mathbb{P}[\text{device 2 } \textbf{operates}] \\
&\times \cdots \times \mathbb{P}[\text{device n } \textbf{operates}]
\end{aligned}
$$

- Parallel devices:

$$
\begin{aligned}
\mathbb{P}[\text{Circuit } \textbf{fails}] = \mathbb{P}[\text{device 1 } \textbf{fails}] \times \mathbb{P}[\text{device 2 } \textbf{fails}] \\
\times \cdots \mathbb{P}[\text{device n } \textbf{fails}]
\end{aligned}
$$

- Remember :

$$\mathbb{P}[\text{failure}] = 1 - \mathbb{P}[\text{success}]$$

5 Techniques of Counting

- Arranging n items in n places : number of ways

$$n!$$

- Permutations :

$$^{n}P_k = \frac{n!}{(n-k)!}$$

- Combinations :

$$^{n}C_k = \frac{n!}{k!\,(n-k)!} = \frac{^{n}P_k}{k!}$$

	order	no order
replacement	n^r	$^{n-r+1}C_r$
no replacement	$^{n}P_r$	$^{n}C_r$

6 Discrete Random Variables

- What are random variables?
  Random variables are mappings from events to numbers.

- probability mass function (PMF) of a random variable X is a function which specifies the probability of obtaining a number x.

$$p_X(x) = \mathbb{P}[X = x]$$

- Note that a PMF should satisfy the following condition

$$\sum_{x \in X(\Omega)} p_X(x) = 1$$

- Cumulative distribution function CDF :

$$F_X(x_k) = \mathbb{P}[X \leq x_k] = \sum_{l=-\infty}^{k} p_X(l)$$

- What is expectation?
  Expectation = Mean = Average computed from a PMF.

$$\mathbb{E}[X] = \mu = \sum_{x \in X(\Omega)} x p_X(x)$$

- Properties:

$$\mathbb{E}[g(X)] = \sum_{x \in X(\Omega)} g(X) p_X(x)$$

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$$

- What is variance?
  It is a measure of the deviation of the random variable X relative to its mean.

$$
\begin{aligned}
\text{Var}[X] = \sigma^2 &= \mathbb{E}[(X-\mu)^2] \\
&= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&= \mathbb{E}[X^2] - \mu^2
\end{aligned}
$$

- Properties:

$$\text{Var}[aX + b] = a^2 \text{Var}[X]$$

- Coefficient of variance $= \dfrac{\sigma}{\mu}$

7 Special Discrete Random Variables

7.1 Bernoulli

(a coin-flip random variable)

- $\mathbb{P}[\text{sucess}] = p$, $\mathbb{P}[\text{failure}] = 1 - p = q$
- PMF :
$$p_X(0) = 1 - p \qquad p_X(1) = p$$
- Expectation:
$$\mathbb{E}[X] = p$$
- Variance:
$$\begin{aligned}\mathrm{Var}[X] &= p(1-p)\\ &= pq\end{aligned}$$

7.2 Bionomial

(n times coin-flips random variable)

- $\mathbb{P}[\text{sucess}] = p$, $\mathbb{P}[\text{failure}] = 1 - p = q$
- PMF :
$$p_X(k) = {}^nC_k p^k q^{n-k}$$
- Expectation:
$$\mathbb{E}[X] = np$$
- Variance:
$$\begin{aligned}\mathrm{Var}[X] &= np(1-p)\\ &= npq\end{aligned}$$
- Show that the binomial PMF sums to 1.:
Use the binomial theorem:
$$\begin{aligned}\sum_{k=0}^{n} p_X(k) &= \sum_{k=0}^{n} {}^nC_k p^k q^{n-k}\\ &= (p + (1-p))^n\\ &= 1\end{aligned}$$

7.3 Geometric

(Trying a binary experiment until we succeed random variable)

- $\mathbb{P}[\text{sucess}] = p$, $\mathbb{P}[\text{failure}] = 1 - p = q$
- PMF :
$$p_X(k) = \underbrace{(1-p)^{k-1}}_{k-1\text{failures}} \underbrace{p}_{\text{final success}}$$
- Expectation:
$$\mathbb{E}[X] = \frac{1}{p}$$
- Variance:
$$\begin{aligned}\mathrm{Var}[X] &= \frac{1-p}{p^2}\\ &= \frac{q}{p^2}\end{aligned}$$

7.4 Poisson

(For small p and large n where $\lambda = np$)

- λ = the rate of the arrival
- PMF :
$$p_X(k) = e^{-\lambda}\frac{\lambda^k}{k!}$$
- Expectation:
$$\mathbb{E}[X] = \lambda$$
- Variance:
$$\mathrm{Var}[X] = \lambda$$
- Show that the Poisson PMF sums to 1.:
Use the exponential series:
$$\begin{aligned}\sum_{k=0}^{\infty} p_X(k) &= \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!}\\ &= e^{-\lambda}\underbrace{\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}}_{=e^{\lambda}}\\ &= 1\end{aligned}$$

8 Continuous Random Variables

- probability density function (PDF) is a continuous version of a PMF, we integrate PDF to compute the probability
$$\mathbb{P}[a \le X \le b] = \int_a^b f_X(x)\,dx$$
- Note that a PMF should satisfy the following condition
$$\int_\Omega f_X(x)\,dx = 1$$
- Note :
$$\mathbb{P}[X = \text{certian point}] = 0$$
- Cumulative distribution function CDF :
$$F_X(x_k) = \mathbb{P}[X \le x] = \int_{-\infty}^{x} f_X(t)dt$$
- Note:
$$\mathrm{CDF} = \int \mathrm{PDF}$$
$$\mathrm{PDF} = \frac{d}{dx}\mathrm{PDF}$$
- Expectation (Mean):
$$\mathbb{E}[X] = \mu = \int_\Omega x\, f_X(x)dx$$
- Properties:
$$\mathbb{E}[g(X)] = \mu = \int_\Omega g(X)\, f_X(x)dx$$
$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$$
- Mode: the peak of the PDF

  How to find mode from PDF:
  • Find a point c such that $f_x(c)$ is maximized by differentiation (and test the edges of the interval).
  How to find mode from CDF:
  • Continuous: Find a point c such that $F_X(c)$ has the steepest slope.
  • Discrete: Find a point c such that $F_X(c)$ has the biggest gap in a jump.
- Median: (a point c that separates the PDF into two equal areas)
$$\mathbb{P}[x < c] = \mathbb{P}[x > c] = 0.5$$
$$F_X(c) = 0.5$$
- Note : Symmetric distribution is a distribution in which Median = Mean
- Percentiles:
To get the α percentile, find the value c at which
$$F_X(c) = \alpha$$
- Variance:
$$\begin{aligned}\mathrm{Var}[X] = \sigma^2 &= \mathbb{E}[(X-\mu)^2]\\ &= \int_\Omega (x-\mu)^2 f_X(x)dx\\ &= \mathbb{E}[X^2] - \mathbb{E}[X]^2\\ &= \mathbb{E}[X^2] - \mu^2\end{aligned}$$
- Properties:
$$\mathrm{Var}[aX + b] = a^2\mathrm{Var}[X]$$

9 Special Continuous Random Variables

9.1 Uniform

- PDF :
$$f_X(x) = \begin{cases}\frac{1}{b-a} & a \le x \le b\\ 0 & \text{otherwise}\end{cases}$$
- Expectation:
$$\mathbb{E}[X] = \frac{a+b}{2}$$
- Variance:
$$\mathrm{Var}[X] = \frac{(a-b)^2}{12}$$

9.2 Exponential

- What is the origin of exponential random variables?
  - An exponential random variable is the *interarrival* time between two consecutive Poisson events.
- PDF :
$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$
- CDF :
$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \geq 0 \end{cases}$$
- Expectation:
$$\mathbb{E}[X] = \frac{1}{\lambda}$$
- Variance:
$$\text{Var}[X] = \frac{1}{\lambda^2}$$
- Memorylessness property:
$$\mathbb{P}[T < t + m | T > t] = \mathbb{P}[T < m] = F_X(m)$$
- Starting from poisson distribution, derive an expression of PDF of exponential random variable

We assume that N is Poisson with a parameter λtor any duration t :

$\mathbb{P}[N = n] = e^{-\lambda t}\dfrac{(\lambda t)^n}{n!}$

Let T be the interarrival time between two events

$\mathbb{P}[T > t] = \mathbb{P}[\text{interarrival time} > t] = \mathbb{P}[\text{no arrival in t}]$

$= \mathbb{P}[N = 0] = e^{-\lambda t}\dfrac{(\lambda t)^0}{0!} = e^{-\lambda t}$

$$\text{since}\mathbb{P}[T > t] = 1 - F_T(t)$$
$$\therefore F_T(t) = 1 - e^{-\lambda t}$$
$$f_T(t) = \frac{d}{dx}F_T(t) = \lambda e^{-\lambda t}$$

9.3 Erlange-k

(A generalization of the exponential distribution is the length until r counts occur in a Poisson process.)

- PDF :
$$f_X(x) = \frac{\lambda^k e^{-\lambda x} x^{k-1}}{(k-1)!}$$
- Expectation:
$$\mathbb{E}[X] = \frac{k}{\lambda}$$
- Variance:
$$\text{Var}[X] = \frac{k}{\lambda^2}$$

9.4 Gamma

- PDF :
$$f_X(x) = \frac{1}{\beta^r \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}$$
α : Shape parameter
β : Scale parameter
- Expectation:
$$\mathbb{E}[X] = \alpha\beta$$
- Variance:
$$\text{Var}[X] = \alpha\beta^2$$
- Starting from gamma distribution, derive an expression of PDF for erlang-k random variable

$$f_X(x) = \frac{1}{\beta^r \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}$$

Substitute $\alpha = k$ and $\beta = \dfrac{1}{\lambda}$

$$f_X(x) = \frac{\lambda^k e^{-\lambda x} x^{k-1}}{\Gamma(k)}$$

If k is an integer, X has an Erlang distribution.

$$f_X(x) = \frac{\lambda^k e^{-\lambda x} x^{k-1}}{(k-1)!}$$

- Exponential distribution is a special case of Gamma distribution with $\alpha = 1$ and $\beta = \dfrac{1}{\lambda}$
- Chi-Squared distribution χ^2 is a special case of Gamma distribution with $\alpha = v/2$ and $\beta = 2$
it is a important distribution in statistics, also called as number of degrees of freedom

9.5 Gaussian

- We write
$$X \sim \mathcal{N}(\mu, \sigma^2)$$
- PDF :
$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$
- Expectation:
$$\mathbb{E}[X] = \mu$$
- Variance:
$$\text{Var}[X] = \sigma^2$$

9.6 Standart Gaussian

- We write
$$Z \sim \mathcal{N}(0, 1)$$
- Conversion from Gaussian to Standard Gaussian
$$Z = \frac{X - \mu}{\sigma}$$
- PDF :
$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\dfrac{z^2}{2}}$$
- CDF:
$$\Phi(z) = \mathbb{P}[Z < z]$$
$$\mathbb{P}[Z > z] = 1 - \Phi(z)$$
$$\Phi(-z) = 1 - \Phi(z)$$
- Expectation:
$$\mathbb{E}[X] = \mu = 0$$
- Variance:
$$\text{Var}[X] = \sigma^2 = 1$$

10 Moment generating function

- MGF:
$$M_X(t) = \mathbb{E}[e^{tX}]$$
- r$^{\text{th}}$ moment:
$$\mathbb{E}[X^r] = \frac{d^r}{dt^r} M_X(t)\Big|_{t=0}$$

11 Joint Discrete Probability Distributions

-
$$f_{XY}(x, y) = \mathbb{P}[X = x, Y = y]$$
- Properties :
$$\sum_X \sum_Y f_{XY}(x, y) = 1$$
- Marginal PMF :
$$f_X(x) = \sum_Y f_{XY}(x, y)$$
$$f_Y(x) = \sum_X f_{XY}(x, y)$$
- Independence : X and Y are independent if
$$\underbrace{f_{XY}(x, y) = f_X(x) \times f_Y(y)}_{\text{for all values of x and y}}$$

also if :
$$f_{X|Y} = f_X$$
$$f_{Y|X} = f_Y$$
- Conditional probability:
$$f_{Y|X} = \frac{f_{XY}(x, y)}{f_X(x)}$$
$$f_{X|Y} = \frac{f_{XY}(x, y)}{f_Y(y)}$$

12 Joint Continuous Probability Distributions

- $$\mathbb{P}[\Lambda] = \iint_{\Lambda} f_{XY}\ dx\ dy$$
  for any event $\Lambda \subseteq \Omega_X \times \Omega_Y$
- Properties :
  $$\iint_{A} f_{XY}(x,\ y)\ dA = 1$$
- Marginal PMF :
  $$f_X(x) = \int_{Y} f_{XY}(x,\ y)\ dy$$
  $$f_Y(x) = \int_{X} f_{XY}(x,\ y)\ dx$$
- Independence : X and Y are independent if
  $$f_{XY}(x,y) = f_X(x) \times f_Y(y)$$
  also if :
  $$f_{X|Y} = f_X$$
  $$f_{Y|X} = f_Y$$
- Conditional probability:
  $$f_{Y|X} = \frac{f_{XY}(x,\ y)}{f_X(x)}$$
  $$f_{X|Y} = \frac{f_{XY}(x,\ y)}{f_Y(y)}$$

13 Expectation, Covariance and Correlation Coefficient

- Discrete:
  $$\mathbb{E}[g(x,\ y)] = \sum_{X}\sum_{Y} g(x,\ y)\, f_{XY}(x,\ y)$$
- Continuous:
  $$\mathbb{E}[g(x,\ y)] = \int_X \int_Y g(x,\ y)\, f_{XY}(x,\ y)dx\, dy$$
- Properties:
  $$\mathbb{E}[x+y] = \mathbb{E}[x] + \mathbb{E}[y]$$
- if x and y are independent:
  $$\mathbb{E}[xy] = \mathbb{E}[x]\mathbb{E}[y]$$
- Covariance:
  $$\begin{aligned}\mathrm{Cov}(X,\ Y) = \sigma_{XY} &= \mathbb{E}[(X-\mu_X)(Y-\mu_Y)] \\ &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]\end{aligned}$$
- Variance:
  $$\mathbb{V}[X+Y] = \mathbb{V}[X] + \mathbb{V}[Y] + 2\mathrm{Cov}(X,\ Y)$$
- if x and y are independent:
  $$\mathbb{V}[X+Y] = \mathbb{V}[X] + \mathbb{V}[Y]$$
- Correlation Coefficient:
  $$\rho_{X,Y} = \frac{\mathrm{Cov}(X,\ Y)}{\sigma_X \sigma_Y}$$

14 Random Processes

- Expectation:
  $$\mu_X(t) = \mathbb{E}[X(t,A)] = \sum_{A} X(t,A) f_A(a)$$
  $$\mu_X(t) = \mathbb{E}[X(t,A)] = \int_A X(t,A) f_A(a)\, dA$$
- Auto-correlation function :
  $$R_{XX}(t, t+\tau) = \mathbb{E}[X(t)X(t+\tau)]$$
- Auto-covariance function :
  $$\begin{aligned}\mathrm{Cov}_{XX}(t,t+\tau) &= R_{XX}(t,t+\tau) - \mu_X(t)\mu_X(t+\tau) \\ &= \mathbb{E}[X(t)X(t+\tau)] - \mathbb{E}[X(t)]\mathbb{E}[X(t+\tau)]\end{aligned}$$
- Wide Sense Stationary Process WSSP:
  $$\text{Expectaion} = \text{Constant, Not depend on time}$$
  $$R_{XX}(t,t+\tau) = \mathbb{E}[X(t)X(t+\tau)] = R_{XX}(\tau)$$
  depend on time difference only
- Average power for WSSP:
  $$R_{XX}(\tau = 0) = \mathbb{E}[X(t)X(t+0)] = \mathbb{E}[X^2(t)]$$