AML5103: Applied Probability & Statistics – Formula Sheet – November 23, 2023

1. **Inclusion–exclusion principle for 4 events**:

$$\begin{aligned} P\left(A_1 \cup A_2 \cup A_3 \cup A_4\right) = {} & P(A_1) + P(A_2) + P(A_3) + P(A_4) \\ & - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_1 \cap A_4) - P(A_2 \cap A_3) - P(A_2 \cap A_4) - P(A_3 \cap A_4) \\ & + P(A_1 \cap A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_4) + P(A_1 \cap A_3 \cap A_4) + P(A_2 \cap A_3 \cap A_4) \\ & - P(A_1 \cap A_2 \cap A_3 \cap A_4). \end{aligned}$$

2. **Conditional probability**:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \Leftrightarrow P(A \cap B) = P(A \mid B) \times P(B).$$

3. **Law of total probability**:
   For event A and the mutually exclusive & collectively exhaustive set of events $\{B_1, B_2, \ldots, B_k\}$ for the same sample space,

$$\begin{aligned} P\Bigg(\underbrace{A}_{\text{e.g., has Cancer}}\Bigg) &= P\Bigg(\underbrace{A \cap B_1}_{\text{has Cancer AND from State-1}}\Bigg) + P\Bigg(\underbrace{A \cap B_2}_{\text{has Cancer AND from State-2}}\Bigg) + \cdots + P\Bigg(\underbrace{A \cap B_k}_{\text{has Cancer AND from State-k}}\Bigg) \\ &= P(A \mid B_1) \times P(B_1) + P(A \mid B_2) \times P(B_2) + \cdots + P(A \mid B_k) \times P(B_k). \end{aligned}$$

   Additional conditioning form:

$$\begin{aligned} P\Bigg(\underbrace{A}_{\text{e.g., has Cancer}} \mid \text{male}\Bigg) &= P\Bigg(\underbrace{A \cap B_1}_{\text{has Cancer AND from State-1}} \mid \text{male}\Bigg) + P\Bigg(\underbrace{A \cap B_2}_{\text{has Cancer AND from State-2}} \mid \text{male}\Bigg) + \cdots + P\Bigg(\underbrace{A \cap B_k}_{\text{has Cancer AND from State-k}} \mid \text{male}\Bigg) \\ &= P(A \mid B_1, \text{male}) \times P(B_1 \mid \text{male}) + P(A \mid B_2, \text{male}) \times P(B_2 \mid \text{male}) + \cdots + P(A \mid B_k, \text{male}) \times P(B_k \mid \text{male}). \end{aligned}$$

4. **Bayes' theorem**:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B \cap A)}{P(B)} = \frac{P(B \mid A) \times P(A)}{P(B)} = \frac{P(B \mid A) \times P(A)}{P(B \mid A) \times P(A) + P(B \mid A^c) \times P(A^c)}$$

   Additional conditioning form:

$$P(A \mid B, C) = \frac{P(B \mid A, C) \times P(A \mid C)}{P(B \mid C)} = \frac{P(B \mid A, C) \times P(A \mid C)}{P(B \mid A, C) \times P(A \mid C) + P(B \mid A^c, C) \times P(A^c \mid C)}.$$

5. **Discrete random variables**:

Random Variable X	**Parameters**	**Values Taken** j or x	**PMF** $P_X(j) = P(X=j)$	**Expected Value** $E[X] = \sum_j jP_X(j)$	**Variance** $Var[X] = \sum_j (j-E[X])^2 P_X(j)$
Bernoulli	p	$j = 0, 1$	$p^j(1-p)^{1-j}$	p	$p(1-p)$
Multinoulli	$p_1, p_2, \ldots, p_k$	$x = \underbrace{\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix}}_{\text{assuming } k=4}$	$\prod_{i=1}^k p_i^{I(x_i=1)}$	$\underbrace{\begin{bmatrix}p_1\\p_2\\p_3\\p_4\end{bmatrix}}_{\text{assuming } k=4}$	$\underbrace{\begin{bmatrix}p_1(1-p_1)\\p_2(1-p_2)\\p_3(1-p_3)\\p_4(1-p_4)\end{bmatrix}}_{\text{assuming } k=4}$
Binomial	n, p	$j = 0, 1, \ldots, n$	$\binom{n}{j}p^j(1-p)^{n-j}$	np	$np(1-p)$
Multinomial	$n, p_1, p_2, \cdots, p_k$	$x = \underbrace{\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}, \text{ where } \sum_{i=1}^4 x_i = n}_{\text{assuming } k=4}$	$\underbrace{\binom{n}{x_1, x_2, x_3, x_4}\prod_{i=1}^4 p_i^{x_i}}_{\text{assuming } k=4}$	$\underbrace{\begin{bmatrix}np_1\\np_2\\np_3\\np_4\end{bmatrix}}_{\text{assuming } k=4}$	$\underbrace{\begin{bmatrix}np_1(1-p_1)\\np_2(1-p_2)\\np_3(1-p_3)\\np_4(1-p_4)\end{bmatrix}}_{\text{assuming } k=4}$
Negative Binomial	r, p	$j = r, r+1, \ldots$	$\binom{j-1}{r-1}p^r(1-p)^{j-r}$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$
Geometric	p	$j = 1, 2, \ldots$	$(1-p)^{j-1}p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Poisson	λ	$j = 0, 1, \ldots$	$\frac{e^{-\lambda}\lambda^j}{j!}$	λ	λ
Hypergeometric	$n_{success}, n_{failure}, n$	$j = 0, 1, \ldots, n$	$\frac{\binom{n_{success}}{j}\binom{n_{failure}}{n-j}}{\binom{n_{success}+n_{failure}}{n}}$	np, where $p = \frac{n_{success}}{n_{success}+n_{failure}}$	$\frac{np(1-p)(N-n)}{N-1}$, where $N = n_{success} + n_{failure}$

6. **Continuous random variables**:

Random Variable	**Parameters**	**PDF**	**Expected Value**	**Variance**
Normal	μ, σ	$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2
Exponential	λ	$f_T(t) = \begin{cases}\lambda e^{-\lambda t} & \text{if } t \geq 0,\\ 0 & \text{if } t < 0.\end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Uniform	a, b	$f_X(x) = \begin{cases}\frac{1}{b-a}, & \text{if } a \leq x \leq b,\\ 0 & \text{otherwise.}\end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$