STA9715
Test 1 - Formula Sheet

Foundations:

- Probabilities are Non-Negative: $0 \leq \mathbb{P}(A) \leq 1$ for all *events* A
- Probability of Entire Sample Space: $\mathbb{P}(\Omega) = 1$
- Probability of Empty Set: $\mathbb{P}(\emptyset) = 0$
- 'Union Bound': $\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$ with equality if A, B are *disjoint* ($A \cap B = \emptyset$)
- Complements: $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
- Naive Probability: $\mathbb{P}(A) = |A|/|\Omega|$
- Counting: ${}_nP_k = \frac{n!}{(n-k)!}$ (permutations - *ordered* choice of k from n); ${}_nC_k = \binom{n}{k} = \frac{n!}{(n-k)!k!}$ (combinations - *unordered*)
- DeMorgan's Laws: $(A \cup B)^c = A^c \cap B^c$ and $(A \cap B)^c = A^c \cup B^c$

Random Variables (Discrete):

- $\mathbb{P}(X = x) = f_X(x)$ is the *probability mass function* of X. Gives the probability of observing $X = x$ exactly
- $\sum_{x \in \mathsf{supp}(X)} \mathbb{P}(X = x) = 1$. $0 \leq f_X(x) \leq 1$

Random Variables (Continuous):

- *Probability density function* of X: $f_X(x)$ Integrates to give the probability X in an interval: $P(a \leq X \leq b) = \int_a^b f_X(x)\,\mathrm{d}x$
- $\int_{x \in \mathsf{supp}(X)} f_X(x)\,\mathrm{d}x = 1. f_X(x)$ may be greater than 1 for small ranges; never negative

Moments:

- Expectation: $\mathbb{E}[X] = \sum x * f_X(x)$ or $\int x * f_X(x)\,\mathrm{d}x$
- Linearity of Expectation: $\mathbb{E}[aX + bY + c] = a\mathbb{E}[X] + b\mathbb{E}[Y] + c$
- Expectation of Functions: $\mathbb{E}[g(X)] = \sum g(x) f_X(x)$ or $\int g(x) f_X(x)\,\mathrm{d}x$
- Variance: $\mathbb{V}[X] = \mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \geq 0$
- $\mathbb{V}[aX + bY + c] = a^2\mathbb{V}[X] + b^2\mathbb{V}[Y]$ only if X, Y are uncorrelated. (Independent implies uncorrelated)
- Expectation of Indicators: $\mathbb{E}[1_{\cdot \in A}(X)] = \mathbb{P}(X \in A)$. Useful to reduce probabilities to linear expectation calculations

Conditional Probabilities:

- $\mathbb{P}(A|B) = \mathbb{P}(A \cap B)/\mathbb{P}(B)$. Special case: $A \subseteq B \implies \mathbb{P}(A|B) = \mathbb{P}(A)/\mathbb{P}(B)$
- Bayes' Rule: $\mathbb{P}(A|B) = \mathbb{P}(B|A) * \mathbb{P}(A)/\mathbb{P}(B) = \mathbb{P}(B|A) * \mathbb{P}(A)/(\mathbb{P}(B|A) * \mathbb{P}(A) + \mathbb{P}(B|A^c) * \mathbb{P}(A))$
- Law of Total Probability: if $\{A_j\}$ are a disjoint partition of Ω then $\mathbb{P}(B) = \sum_j \mathbb{P}(B|A_j)\mathbb{P}(A_j)$
- Law of Total Expectation: $\mathbb{E}[X] = \sum \mathbb{E}[X|A_i]\mathbb{P}(A_i)$ for partition $\{A_i\}$ or $\mathbb{E}_X[X] = \mathbb{E}_Y[\mathbb{E}_X[X|Y]]$
- Law of Total Variance: $\mathbb{V}[X] = \mathbb{E}_Y[\mathbb{V}_X[X|Y]] + \mathbb{V}_Y[\mathbb{E}_X[X|Y]]$
- Independence: A, B are *independent* if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Equivalently $\mathbb{P}(B|A) = \mathbb{P}(B)$ and $\mathbb{P}(A|B) = \mathbb{P}(A)$

Distributions:

- Bernoulli: $X \sim \mathrm{Bern}(p)$. $\mathbb{P}(X = 1) = p; \mathbb{P}(X = 0) = 1 - p$. $\mathbb{E}[X] = p$; $\mathbb{V}[X] = p(1 - p)$. 'Coin Flip'
- Binomial: Sum of n IID Bernoulli: $P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$. $\mathbb{E}[X] = np$. $\mathbb{V}[X] = np(1 - p)$
- Poisson: Limit of $n \to \infty, p \to 0$ Binomial. $X \sim \mathrm{Pois}(\mu)$. $\mathbb{P}(X = x) = \mu^x e^{-\mu}/x!$. $\mathbb{E}[X] = \mathbb{V}[X] = \mu$
- Geometric: IID Bernoulli 'until' 1st success: $X \sim \mathrm{Geom}(p)$. $\mathbb{P}(X = x) = p(1 - p)^{x-1}$. $\mathbb{E}[X] = 1/p$. $\mathbb{V}[X] = (1 - p)/p^2$. Memoryless
- Normal: $X \sim \mathcal{N}(\mu, \sigma^2)$. $f_X(x) = e^{-(x-\mu)^2/2\sigma^2}/\sqrt{2\pi\sigma^2}$. $\mathbb{E}[X] = \mu$. $\mathbb{V}[X] = \sigma^2$. 'Bell curve'
- Exponential: $X \sim \mathrm{Exp}(\lambda)$. $f_X(x) = \lambda e^{-\lambda x}$. $\mathbb{E}[X] = \lambda^{-1}$. $\mathbb{V}[X] = \lambda^{-2}$. Continuous geometric
- Uniform (Discrete and Continuous). $\mathbb{E}[\mathrm{DUnif}\{a, \dots, b\}] = \mathbb{E}[\mathrm{CUnif}([a, b])] = (a + b)/2$. $\mathbb{V}[\mathrm{CUnif}([a, b])] = (b - a)^2/12$