Looking for comprehensive study materials for the Capstone Exam? Explore our:
**All-In-One Capstone Exam Review and Reference Materials**

## A1.1 PROBABILITY

**Conditional Probability of Event $B$ Given Event $A$**: $P[B|A] = \dfrac{P[B \cap A]}{P[A]}$

**Probability of Event $B$ Partitioned by Events $A$ and $A'$**:

$$P[B] = P[B \cap A] + P[B \cap A'] = P[B|A] \cdot P[A] + P[B|A'] \cdot P[A']$$

**Bayes' Rule**: $P[A|B] = \dfrac{P[A \cap B]}{P[B]} = \dfrac{P[B|A] \cdot P[A]}{P[B|A] \cdot P[A] + P[B|A'] \cdot P[A']}$

**Independent Events $A$ and $B$**: $P(A \cap B) = P(A) \cdot P(B)$, $P(A|B) = P(A)$ or $P(B|A) = P(B)$.

**Mutually Independent Events**: $P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{i=1}^{k} P(A_{i_j})$.

**Factorial Notation**: $n! := n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1$

0! is defined to be equal to 1.

**Permutations**: The number of ways of choosing an ordered subset of size $k$ **without replacement** from a collection of $n$ objects:

$$\frac{n!}{(n-k)!} = n \times (n-1) \times \cdots \times (n-k+1)$$

and is denoted by ${}_nP_k$ or $P_{n,k}$ or $P(n,k)$.

**Combinations**: the number of ways of choosing a subset of size $k \leq n$ without replacement and without regard to the order in which the objects are chosen:

$$\frac{n!}{k! \cdot (n-k)!},$$

which is usually denoted by $\binom{n}{k}$, or ${}_nC_k$, $C_{n,k}$, $C(n,k)$, and is read "$n$ choose $k$".

**Binomial Coefficient - Number of Subsets of Size $k$ From a Collection of $n$ Objects**: $\binom{n}{k} = \dfrac{n!}{k! \cdot (n-k)!}$

**DeMorgan's Laws**: $(A \cup B)' = A' \cap B'$ and $(A \cap B)' = A' \cup B'$.

**Inclusion-Exclusion Equations**:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P[A \cup B \cup C] = P[A] + P[B] + P[C] - P[A \cap B] - P[B \cap C] - P[A \cap C] + P[A \cap B \cap C]$$

**Central Limit Theorem**: Suppose that $X$ is a random variable with mean $\mu$ and standard deviation $\sigma$ and suppose that $X_1, X_2, \cdots, X_n$ are $n$ independent random variables with the same distribution as $X$. Let $Y_n = X_1 + X_2 + \cdots + X_n$.

Then $E[Y_n] = n\mu$ and $Var[Y_n] = n\sigma^2$, and as $n$ increases, the distribution of $Y_n$ approaches a normal distribution $N(n\mu, n\sigma^2)$.

## A1.2 UNIVARIATE RANDOM VARIABLES

**Probability Mass Function of a Discrete Random Variable**:

Denoted $p(x)$, $p_X(x)$ or $p_x$, equal to the probability that the value $x$ occurs, $P(X = x)$.

**Probability Density Function of a Continuous Random Variable**:

Denoted $f(x)$ or $f_X(x)$, and it must be non-negative ($f(x) \geq 0$) on the intervals on which $X$ is defined.

The probability that $X$ is in the interval $(a, b)$ is defined as

$$P(a < X < b) := \int_a^b f(x)dx.$$

**Distribution Function & Survival Function of Random Variable $X$**:

$F_X(t) = P[X \leq t]$ $\qquad$ $S_X(t) = 1 - F_X(t) = P[X > t]$

$F(x) = \int_{-\infty}^x f(t)dt$ $\qquad$ $\frac{d}{dx}F(x) = F(x) = -S'(x) = f(x)$

**Hazard Rate Function**: $h(x) = \dfrac{f(x)}{1 - F(x)} = -\dfrac{d}{dx}\log(1 - F(x))$

**Conditional distribution of $X$ given event $A$**:

$$f_{X|A}(x|A) = \begin{cases} \dfrac{f(x)}{P(A)} & \text{if } x \text{ is an outcome in event } A \\ 0 & \text{if } x \text{ is not an outcome in event } A \end{cases}$$

**Expected Value of Random Variable $X$**:

$E[X] = \sum_{\text{all } x_i} x_i \cdot p_X(x_i)$ (discrete)

$E[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)\, dx$ (continuous)

**Expectation of $h(x)$**: If $h$ is a function, then

$$\mathrm{E}[h(X)] = \begin{cases} \sum_x h(x)p(x) & \text{if } X \text{ is a discrete random variable} \\ \int_{-\infty}^{\infty} h(x)f(x) & \text{if } X \text{ is a continuous random variable} \end{cases}$$

For any constants $a_1, a_2$ and $b$ and functions $h_1$ and $h_2$,

$$\mathrm{E}[a_1h_1(X) + a_2h_2(X) + b] = a_1\mathrm{E}[h_1(X)] + a_2\mathrm{E}[h_2(X)] + b$$

As a special case, $\mathrm{E}[aX + b] = a\mathrm{E}[X] + b$.

**Moment of Random Variable $X$**: If $n \geq 1$ is an integer, the $\boldsymbol{n}$-th moment of $X$ is $E[X^n]$.

If the mean of $X$ is $\mu$, then the $\boldsymbol{n}$-th central moment of $X$ is $E[(X - \mu)^n]$.

**Variance**: $\mathrm{Var}[X] = \mathrm{E}[(X - \mu_x)^2] = \mathrm{E}[(X - \mathrm{E}[X])^2]$

$\mathrm{Var}[X] = \mathrm{E}[X^2] - (\mathrm{E}[X])^2$

If $a$ and $b$ are constants, then $\mathbf{Var}(aX + b) = a^2\mathbf{Var}(X)$.

**Standard Deviation**: $\sigma_X = \sqrt{\text{Var}(X)}$.

**Coefficient of Variation of $X$ :** $\dfrac{\sigma_X}{\mu_X}$ or $\dfrac{\sigma_X}{E[X]}$

**Percentiles of a Distribution**: If $0 < p < 1$, then the **100$p$-th percentile** of the distribution of $X$ is the number $c_p$ which satisfies both of the following inequalities:

$$P(X \leq c_p) \geq p \text{ and } P(X \geq c_p) \geq 1 - p.$$

**Median of the Distribution of Random Variable $X$**:

Smallest $m$ for which $F_X(m) = .5$

**Mode of the Distribution of Random Variable $X$**:

The point $c$ at which $p_X(c)$ (discrete) or $f_X(c)$ (continuous) is maximized.

**Discrete Uniform Distribution on the Integers $1, 2, \ldots, N$**:

$p_X(k) = \dfrac{1}{N}$ for $k = 1, 2, \ldots N$  $\quad$ $\text{E}[X] = \dfrac{N+1}{2}$ $\quad$ $\textbf{Var}[X] = \dfrac{N^2 - 1}{12}$

**Binomial Distribution for $n$ Trials and Success Probability $p$**:

$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$ $\quad$ $\text{E}[X] = np$ $\quad$ $\textbf{Var}[X] = np(1-p)$

**Geometric Distribution on the Integers $0, 1, 2, \ldots$ and Parameter $p$**:

$p_X(k) = (1-p)^k \cdot p$ $\quad$ $\text{E}[X] = \dfrac{1-p}{p}$ $\quad$ $\textbf{Var}[X] = \dfrac{1-p}{p^2}$

**Poisson Distribution with Mean $\lambda$**:

$p_X(k) = \dfrac{e^{-\lambda}\lambda^k}{k!}$ for $k = 0, 1, 2, \ldots$ $\quad$ $\text{E}[X] = \lambda$ $\quad$ $\textbf{Var}[X] = \lambda$

**Negative Binomial Distribution**: Two Parametrizations

*Successes-Probability Form (Parameters $p$ and integer $\gamma \geq 1$):*

$p_X(x) = \binom{\gamma + x - 1}{x} p^\gamma (1-p)^x$, $x = 0, 1, 2, 3, \ldots$ $\quad$ $\text{E}[X] = \dfrac{\gamma(1-p)}{p}$ $\quad$ $\textbf{Var}[X] = \dfrac{\gamma(1-p)}{p^2}$

*Mean-Dispersion Form (Parameters $\mu$ and integer $\gamma \geq 1$, as in R):*

$p_X(x) = \binom{\gamma + x - 1}{x} \left(\dfrac{\gamma}{\gamma + \mu}\right)^\gamma \left(\dfrac{\mu}{\gamma + \mu}\right)^x$, $\quad$ $\text{E}[X] = \mu$ $\quad$ $\textbf{Var}[X] = \mu + \dfrac{\mu^2}{\gamma}$

$x = 0, 1, 2, 3, \ldots$

*Conversion*: $\mu = \dfrac{\gamma(1-p)}{p}$, $p = \dfrac{\gamma}{\gamma + \mu}$

*Note*: Use Mean-Dispersion form for R simulations (`rnbinom(n, size=`$\gamma$`, mu=`$\mu$`)`).

**Hypergeometric Distribution**:

$p_X(x) = \dfrac{\binom{K}{x}\binom{M-K}{n-x}}{\binom{M}{n}}$ $\quad$ $\text{E}[X] = \dfrac{nK}{M}$ $\quad$ $\textbf{Var}[X] = \dfrac{nK(M-K)(M-n)}{M^2(M-1)}$

for $\max\{0, n - (M - K)\} \leq x \leq \min\{n, K\}$

**Continuous Uniform Distribution on the Interval $[a, b]$:**

$f_X(x) = \frac{1}{b-a}$ for $a \le x \le b$ $\quad$ $E[X] = \frac{a+b}{2}$ $\quad$ $\text{Var}[X] = \frac{(b-a)^2}{12}$

**Standard Normal Distribution $N(0,1)$ (Mean 0, Variance 1):**

$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ for $-\infty < x < \infty$ $\quad$ $F_X(x) = \Phi(x)$

**Normal Distribution $N(\mu, \sigma)$ (Mean $\mu$, Variance $\sigma^2$):**

$\phi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$ for $-\infty < x < \infty$ $\quad$ $F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$

**Exponential Distribution**: Two Parametrizations

*Mean Form (Parameter $\theta > 0$):*

$f_X(x) = \frac{1}{\theta} e^{-x/\theta}$, $x > 0$, and $f_X(x) = 0$ otherwise $\quad$ $E[X] = \theta$ $\quad$ $\text{Var}[X] = \theta^2$

*Rate Form (Parameter $\lambda > 0$):*

$f_X(x) = \lambda e^{-\lambda x}$, $x > 0$, and $f_X(x) = 0$ otherwise $\quad$ $E[X] = \frac{1}{\lambda}$ $\quad$ $\text{Var}[X] = \frac{1}{\lambda^2}$

*Conversion*: $\theta = \frac{1}{\lambda}$, $\lambda = \frac{1}{\theta}$

**Gamma Distribution**: Two Parametrizations

*Shape-Scale Form (Parameters $\alpha > 0$, $\theta > 0$):*

$f_X(x) = \frac{x^{\alpha-1} e^{-x/\theta}}{\theta^\alpha \Gamma(\alpha)}$, $x > 0$, and $f_X(x) = 0$ otherwise $\quad$ $E[X] = \alpha\theta$ $\quad$ $\text{Var}[X] = \alpha\theta^2$

*Shape-Rate Form (Parameters $\alpha > 0$, $\beta > 0$):*

$f_X(x) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}$, $x > 0$, and $f_X(x) = 0$ otherwise $\quad$ $E[X] = \frac{\alpha}{\beta}$ $\quad$ $\text{Var}[X] = \frac{\alpha}{\beta^2}$

*Conversion*: $\theta = \frac{1}{\beta}$, $\beta = \frac{1}{\theta}$

**Lognormal Distribution with parameters $\mu$ and $\sigma$:** If $\ln(Y) \sim N(\mu, \sigma^2)$ then we have

$f_Y = \frac{1}{\sigma y \sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\ln(y)-\mu}{\sigma}\right)^2\right)$, $\quad$ $E[Y] = e^{\mu + \frac{\sigma^2}{2}}$ $\quad$ $\text{Var}[Y] = E[Y]^2\left(e^{\sigma^2} - 1\right)$.

**Beta Distribution with parameters $\alpha > 0$ and $\beta > 0$:**

$f_X(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$, $\quad$ $E[X] = \frac{\alpha}{\alpha+\beta}$, $\quad$ $\text{Var}[X] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$.

for $0 < x < 1$ and 0 otherwise.

**Moment Generating Function (MGF):** $\quad$ $M_X(t) = E\left[e^{tX}\right]$

**Probability Generating Function (PGF):** $\quad$ $P_X(t) = E\left[t^X\right]$

If $X$ is **discrete** then $\quad$ $M_X(t) = \sum e^{tx} \cdot p(x)$, $\quad$ $P_X(t) = \sum t^x \cdot p(x)$.

If $X$ is **continuous** then $\quad$ $M_X(t) = \int e^{tx} \cdot p(x)dx$, $\quad$ $P_X(t) = \int t^x \cdot p(x)dx$.

The moments of $X$ can be found from the successive derivatives of $M_X(t)$:

$$M'_X(0) = \mathrm{E}[X],$$
$$M''_X(0) = \mathrm{E}[X^2],$$
$$M_X^{(n)}(0) = \mathrm{E}\left[X^n\right],$$
$$\frac{d^2}{dt^2}\log[M_X(t)]\Big|_{t=0} = \mathrm{Var}(X)$$

If $X$ has a discrete non-negative integer distribution with $p_k = P[X = k]$ and the probability generating function $P_X(t)$, then

$$p_k = \frac{1}{k!}\cdot P_X^{(k)}(0)$$
$$P'_X(1) = \int_{-\infty}^{\infty} x f(x)\,dx = \mathrm{E}[X]$$
$$P''_X(1) = \mathrm{E}[X^2 - X] = \mathrm{E}[X^2] - \mathrm{E}[X]$$

## A1.3 MULTIVARIATE RANDOM VARIABLES

**Cumulative Distribution Function**:

$$F(x,y) = \sum_{s=-\infty}^{x}\sum_{t=-\infty}^{y} f(s,t) \text{ (discrete).}$$
$$F(x,y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f(s,t)dsdt \text{ (continuous).}$$

**Marginal Distribution of $X$ From Joint Distribution of $X$ and $Y$:**

$$p_X(x) = \sum_{\text{all } y} p(x,y) \text{ (discrete)}$$
$$f_X(x) = \int_y f(x,y)dy \text{ (continuous)}$$

**Independence of Random Variables $X$ and $Y$:** $f(x,y) = f_X(x)\cdot f_Y(y), \quad F(x,y) = F_X(x)\cdot F_Y(y)$

**Conditional Distribution of $X$ Given $Y$:** $f_{X|Y}(x|y) = \dfrac{f(x,y)}{f_Y(y)}$

**Conditional Expectation**: $\mathrm{E}[Y|X=x] = \int y f_{Y|X}(y|X=x)\,dy$

**If $X$ and $Y$ are independent random variables**: $f_{Y|X}(y|X=x) = \frac{f(x,y)}{f_X(x)} = \dfrac{f_X(x)f_Y(y)}{f_X(x)} = f_Y(y)$

**Expectation of a Function of Jointly Distributed Random Variables**:

$$\mathrm{E}[h(X,Y)] = \sum_x\sum_y h(x,y)f(x,y) \text{ (discrete)}$$
$$\mathrm{E}[h(X,Y)] = \int_x\int_y h(x,y)f(x,y)dxdy \textbf{ (Continuous)}$$

**Covariance and Correlation Between $X$ and $Y$:**

$$\mathrm{Cov}[X,Y] = \mathrm{E}[(X-\mathrm{E}[X])(Y-\mathrm{E}[Y])] = \mathrm{E}[XY] - \mathrm{E}[X]\cdot\mathrm{E}[Y]$$
$$\rho_{XY} = \frac{\mathrm{Cov}[X,Y]}{\sqrt{\mathrm{Var}[X]\cdot\mathrm{Var}[Y]}}$$

**Variance of the Sum of $X$ and $Y$:** $\mathrm{Var}[aX+bY+c] = a^2\mathrm{Var}[X] + b^2\mathrm{Var}[Y] + 2ab\mathrm{Cov}[X,Y]$

**Joint Moment Generating Function**:

$$M_{X,Y}(t_1,t_2) = \mathrm{E}[e^{t_1X+t_2Y}]$$
$$\mathrm{E}[X^nY^m] = \frac{\partial^{n+m}}{\partial^n t_1\,\partial^m t_2} M_{X,Y}(t_1,t_2)\Big|_{t_1=t_2=0}$$

**Bivariate Normal Distribution**:

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \cdot \exp\left[-\frac{1}{2(1-\rho^2)} \cdot \left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)\right]\right]$$

$$\begin{aligned} \mathrm{E}[Y|X=x] &= \mu_Y + \rho_{XY} \cdot \frac{\sigma_Y}{\sigma_X} \cdot (x-\mu_X) \\ &= \mu_Y + \frac{\mathbf{Cov}(X,Y)}{\sigma_X^2} \cdot (x-\mu_X) \\ \mathbf{Var}(Y|X=x) &= \sigma_Y^2 \cdot (1-\rho_{XY}^2) \end{aligned}$$

**Law of Total Expectation**: $\mathrm{E}[\mathrm{E}[X|Y]] = \mathrm{E}[X]$

**Law of Total Variance**: $\mathrm{Var}(X) = \mathrm{E}[\mathrm{Var}(X|Y)] + \mathrm{Var}(\mathrm{E}[X|Y])$

**Order Statistics: $Y_k$ is the $k$-th from the smallest of the $X_i$'s**

PDF of $Y_k$ is

$$g_k(t) = \frac{n!}{(k-1)!(n-k)!} \cdot (F(t))^{k-1} \cdot (1-F(t))^{n-k} \cdot f(t).$$