MIDTERM STUDY GUIDE

1. Definitions

You should be familiar with the following definitions:

Definition (combination)**.** *The number of unordered subsets of size r from a set of n elements:*

$$C_r^n = \frac{n!}{r!(n-r)!}$$

Definition (compliment)**.** *The compliment of a set A is the set of all elements that do not belong to A:*

$$A^c = \{x : x \notin A\}$$

Definition (conditional probability)**.** *The conditional probability $P(A|B)$ of an event A given B is:*

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Definition (countable set)**.** *A countable set is a set that can be put in 1-1 correspondence with the set of natural numbers $\mathbb{N}$.*

Definition (cumulative distribution function)**.** *The cumulative distribution function (CDF) $F(x)$ of a random variable X is defined as:*

$$F(x) = P(X \leq x)$$

Definition (DeMorgan laws)**.** *If A and B are sets,*

$$(A \cup B)^c = A^c \cap B^c \quad \textit{and} \quad (A \cap B)^c = A^c \cup B^c$$

Definition (discrete probability space)**.** *A discrete probability space is a probability space whose sample space Ω is finite or countably infinite.*

Definition (event space)**.** *The event space $\mathcal{F}$ of an experiment is the collection of subsets of Ω that will be considered to be events.*

Definition (expected value)**.** *The expected value $E(X)$ of a random variable X is defined as:*

$$E(X) = \sum x \cdot p(x) \textit{ if } X \textit{ is discrete or } \int x \cdot f(x)\, dx \textit{ if } X \textit{ is continuous}$$

Definition (expected value of a function)**.** *The the expected value $E(f(X))$ of a function of a random variable X is defined as:*

$$E(f(X)) = \sum f(x){\cdot}p(x) \text{ if } X \text{ is discrete or } \int f(x){\cdot}f(x)\,dx \text{ if } X \text{ is continuous}$$

Definition (experiment)**.** *A process by which an observation is made.*

Definition (independent events)**.** *Two events A and B are independent if any of the following are true:*

$$P(A|B) = P(A) \quad or \quad P(B|A) = B \quad or \quad P(A \cap B) = P(A) \cdot P(B)$$

Definition (Kolmogorov axioms)**.** *The Kologorov axioms are a set of three rules for assigning probabilities to events :*

- $P(E) \geq 0$ *for any* $E \subseteq \Omega$
- $P(\Omega) = 1$
- *If* (P_i) *is a collection of disjoint events,*

$$P(A_1 \cup A_2 \cup A_3 \cdots) = \sum_{i=1}^{\infty} P(A_i)$$

Definition (kth moment)**.** *The kth moment of a random variable X is defined as:*

$$\mu'_k = E(X^k)$$

Definition (kth moment about the mean)**.** *The kth moment about the mean of a random variable X is defined as:*

$$\mu_k = E[(X - \mu)^k]$$

Definition (m times n rule)**.** *If set A has m elements and set B has n elements, there are $m \cdot n$ sets consisting of a signle element from A and a single element from B.*

Definition (moment generating function)**.** *The moment generating function of a random variable X is defined as:*

$$m(t) = E(e^{tx})$$

Definition (outcome)**.** *The result of performing an experiment. Every time an experiment is performed, the result is exactly one outcome or sample point.*

Definition (partition)**.** *A partition P of a set S is a collection of disjoint subsets $\{B_i\}$ whose union is S.*

Definition (permutation). *The number of ordered subsets of size r from a set of n elements:*

$$P_r^n = \frac{n!}{(n-r)!}$$

Definition (pth quantile). *The p^{th} quantile ϕ_p of a random variable X is defined as then mumber ϕ_p with the property:*

$$P(X \leq \phi_p) = p$$

Definition (power set). *The power set $\mathcal{P}$ of a set S is the collection of all possible subsets of S.*

Definition (probability density function). *The probability density function (PDF) $f(x)$ of a continuous random variable X is the derivative of the cumulative distribution function of X.*

Definition (probability mass function). *The probability mass function p of a discrete random variable X maps the values of the random variable into the probability that the random variable takes that value:*

$$p(x) = P(X = x)$$

Definition (probability measure). *A probability measure ρ is a function that maps the event space $\mathcal{F}$ of an experiment into $[0, 1]$:*

$$\rho : \mathcal{F} \to [0, 1]$$

Definition (probability space). *A probability space is a triple $(\Omega, \mathcal{F}, \rho)$ consisting of:*

- *A sample space Ω*
- *An event space $\mathcal{F}$ defined on Ω*
- *A probability measure ρ defined on $\mathcal{F}$*

Definition (random sample). *A random sample is a subset of a population chosen in a way that gives every possible subset an equal chance of being chosen.*

Definition (random variable). *A random variable X is a real-valued function whose domain is a sample space Ω:*

$$X : \Omega \to \mathbb{R}$$

Definition (sample space). *The sample space Ω of an experiment is the set of all possible outcomes or sample points of an experiment.*

Definition (simple event). *A simple event is an event that consists of a single outcome or sample point.*

Definition (standard deviation). *The standard deviation σ of a random variable X is the square root of $V(X)$.*

Definition (uncountable set). *An uncountable set is a set that cannot be put in 1-1 correspondence with the set of natural numbers $\mathbb{N}$.*

Definition (variance). *The variance $V(X)$ of a random variable X is defined as:*

$$V(X) = E[X - E(X)]^2$$

2. Theorems

You should be familiar with the following theorems and know how to apply them.

Theorem (additive law of probability). *If A and B are events,*

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

Theorem (multiplicative law of probability). *If A and B are events,*

$$P(A \cap B) = P(B|A) \cdot P(A) = P(A|B) \cdot P(B)$$

Theorem (multiplicative law of probability - independent events). *If A and B are independent events,*

$$P(A \cap B) = P(A) \cdot P(B)$$

Theorem (probability of a compliment). *If A is an event and A^c is its compliment,*

$$P(A) = 1 - P(A^c)$$

Theorem (Markov's theorem). *If X is a random variable with finite expectation $E(X)$, then*

$$P(|X| \geq a) \leq \frac{E(|X|)}{a} \quad \textit{for any } a > 0$$

Theorem (Markov's theorem - generalized version). *If X is a random variable with finite expectation $E(X)$ and h is a nonnegative real valued function, then*

$$P(h(X) \geq a) \leq \frac{E(h(X))}{a} \quad \textit{for any } a > 0$$

Theorem (Chebychev's theorem). *If X is a random variable with $E(X) = \mu$ and finite variance σ^2, then*

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2} \quad \textit{for any } k > 0$$

Theorem (law of total probability). *If $P = \{B_1, B_2, \ldots, B_n\}$ is partition of a sample space Ω with $P(B_i) > 0$ for each i and A is an arbitrary event,*

$$P(A) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)$$

Theorem (Baye's rule). *If $P = \{B_1, B_2, \ldots, B_n\}$ is partition of a sample space Ω with $P(B_i) > 0$ for each i and A is an arbitrary event,*

$$P(B_j|A) = \frac{P(A|B_j) \cdot P(B_j)}{\sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)}$$

Theorem (moment generating function theorem). *If $m(t)$ is the moment generating function of a random variable X, then the k^{th} moment μ_x' of X is*

$$\mu_k' = \left. \frac{d^k m(t)}{dt^k} \right|_{t=0}$$

Theorem (moment generating function of a sum). *If X and Y are independent random variables with respective moment generating functions $m_x(t)$ and $m_y(t)$, then the moment generating function of the random variable $X + Y$ is*

$$m_{x+y}(t) = m_x(t) \cdot m_y(t)$$

3. Distributions

A copy of the information inside the back cover of the text on continuous distributions and discrete distributions will be provided.

You should be familiar with the following distributions:

3.1. **Bernoulli distribution.**

The Bernoulli distribution arises from an experiment with two outcomes:

- Success S with probability p
- Failure F with probability $1 - p$

The random variable that maps the outcome S to 1 and F to 0 is said to have a Bernoulli distribution.

3.2. **Binomial distribution.**

A binomial random variable arises as a sum of a predetermined number n of independent Bernoulli random variables each with probability of success p.

3.3. **Geometric distribution.** A geometric random variable arises if we conduct independent Bernoulli trials each with probability p until the first success is obtained. The random variable may be defined as either the number of trials Y that this takes (Wackerly text definition) or the number of failures X that precede the first success (R or actuarial exam).

3.4. **Negative binomial distribution.** A negative binomial random variable arises if we conduct independent Bernoulli trials each with probability p until r successes are obtained. The random variable may be defined as either the number of trials Y that this takes (Wackerly text definition) or the number of failures X that precede the r^{th} success (R or actuarial exam).

3.5. **Poisson distribution.** The Poisson distribution arises as a limit of the binomial distribution as $n \to \infty$ while the expected number of successes $np = \lambda$ is held constant.

3.6. **Hypergeometric distribution.** A hypergeometric distribution models the experiment of drawing n chips from an urn containing N chips of which r are red and $N - r$ are black. The random variable is the number of red chips drawn.

3.7. **Uniform distribution.** The uniform distribution arises from the experiment of selecting a number at random from the interval $(0, 1)$ or, more generally, from (θ_1, θ_2) with $\theta_2 > \theta_1$ (It makes no difference whether we consider the endpoints to be included or not).

3.8. **Normal distribution.** The normal distribution is the limiting distribution of a sum of independent random variables as the number of variables increases without bound.

3.9. **Exponential distribution.** The exponential distribution is commonly used to model time to failure of a component or time to some arbitrary event.

3.10. **Gamma distribution.** The gamma distribution can be thought of as a generalization of the exponential.

3.11. **Chi-square distribution.** The chi-square distribution is a special case of the gamma distribution. The square of a standard normal random variable has a chi-square distribution.

3.12. **Beta distribution.** The beta distribution is something like a continuous version of the binomial. It is very flexible and often used in Bayesian statistics to describe the distribution of a parameter.