Lecture
Exam 2 Review

STAT 225 Introduction to Probability Models
March 5, 2014

Agenda

1 Discrete Random Variables

2 Choosing a Discrete Distribution

3 Approximation for Discrete Distributions

4 Nested Problems
**Bernoulli random variable**

Characteristics of the Bernoulli random variable:
Let $X$ be a Bernoulli r.v.

- **The definition of $X$:** It is the number of success in a single trial of a random experiment
- **The support (possible values for $X$):** 0: "failure" or 1: "success"
- **Its parameter(s) and definition(s):** $p$: the probability of success on 1 trial
- **The probability mass function (pmf):**
  \[ p_X(x) = p^x(1-p)^{1-x}, \quad x = 0, 1 \]
- **The expected value:** $E[X] = p$
- **The variance:** $Var(X) = p(1-p)$

---

**Binomial random variable**

Characteristics of the Binomial random variable:
Let $X$ be a Binomial r.v.

- **The definition of $X$:** It is the number of successes in $n$ trials of a random experiment, where sampling is done with replacement (or trials are independent)
- **The support:** $0, 1, \ldots, n$
- **Its parameter(s) and definition(s):** $p$: the probability of success on 1 trial and $n$ is the sample size
- **The probability mass function (pmf):**
  \[ p_X(x) = \binom{n}{x} p^x(1-p)^{n-x}, \quad x = 0, 1, \ldots, n \]
- **The expected value:** $E[X] = np$
- **The variance:** $Var(X) = np(1-p)$
**Hypergeometric random variable:**

Characteristics of the Hypergeometric random variable:
Let $X$ be a hypergeometric r.v.

- The definition of $X$: It is the number of successes in $n$ trials of a random experiment, where sampling is done without replacement (or trials are dependent).
- The support: $k \in \{\max(0, n+K-N), \ldots, \min(n, K)\}$
- Its parameter(s) and definition(s): $N$: the population size, $n$: the sample size, and $K$: number of success in the population
- The probability mass function (pmf):
  $$p_X(x) = \frac{\binom{K}{x} \binom{N-K}{n-x}}{\binom{N}{n}}$$
- The expected value: $E[X] = n\frac{K}{N}$
- The variance: $Var(X) = n\frac{K}{N} \frac{N-K}{N} \frac{N-n}{N-1}$

**Poisson random variable:**

Characteristics of the Poisson random variable:
Let $X$ be a Poisson r.v.

- The definition of $X$: The number of successes per [time interval, area]
- The support: $x = 0, 1, 2, \cdots$
- Its parameter(s) and definition(s): $\lambda$: the average number of successes per [time interval, area]
- The probability mass function (pmf): $p_X(x) = \frac{e^{-\lambda} \lambda^x}{x!}$
- The expected value: $E[X] = \lambda$
- The variance: $Var(X) = \lambda$
**Geometric random variable:**

Characteristics of the Geometric random variable:

Let $X$ be a Geometric r.v.

- The definition of $X$: The number of trials it takes to get the 1st success
- The support: $x = 1, 2, \ldots$
- Its parameter(s) and definition(s): $p$: the probability of success in a single trial
- The probability mass function (pmf):
  \[ p_X(x) = p(1-p)^{x-1} \text{ for } x = 1, 2, \ldots \]
- The expected value: $E[X] = \frac{1}{p}$
- The variance: $\text{Var}(X) = \frac{1-p}{p^2}$

**Properties of Geometric distribution**

- Tail Probability: $P(X > x) = (1-p)^x$
- Memoryless Property:
  \[ P(X > s + t | X > s) = P(X > t) \]

**Negative Binomial random variable:**

Characteristics of the Negative Binomial random variable:

Let $X$ be a Negative Binomial r.v.

- The definition of $X$: The number of trials it takes to get the $r$th success
- The support: $x = r, r+1, r+2, \ldots$
- Its parameter(s) and definition(s): $r$: the number of success of interest, $p$: the probability of success in a single trial
- The probability mass function (pmf):
  \[ p_X(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r} \text{ for } x = r, r+1, r+2, \ldots \]
- The expected value: $E[X] = \frac{r}{p}$
- The variance: $\text{Var}(X) = \frac{rf}{p^2}$
Choosing a Discrete Distribution

- **Bernoulli**
- **Binomial**
- **Hypergeometric**
- **Poisson**
- **Geometric**
- **Negative Binomial**

Approximation for Discrete Distributions

- **Binomial approximation for Hypergeometric**
  
  If $X \sim \text{Hyp}(N, n, K)$ and $N$ is sufficiently larger than $n$, say $N > 20n$, then we can approximate the distribution $X$ by using $X^* \sim \text{Bin}(n, p)$, where $p = \frac{K}{N}$

- **Poisson approximation to Binomial**

  If $X \sim \text{Binomial}(n, p)$ with $n > 100$ and $p < .01$ then we can approximate the distribution $X$ by using $X^* \sim \text{Poisson}(\lambda = n \times p)$
Nested Problems

In some of our examples, we will use a distribution to calculate a probability, then use that probability as the parameter in a new distribution. Take a look of Example 42, 43 in lecture 16 and practice homework set 2 problem 17, 19.