**Probability Basics: Handout 1**

**Math/Stat 394: Probability I**

Wellner; 1/5/2000

**Terminology and Notation:**

**Experiment**

Sample space = Outcome set = $\Omega$;

Events, $A,\ B,\ C,\ \ldots$;

Elementary Outcomes, $a, b, c$, or $i, j, k$

Random Variables, $X, Y, Z$;

Outcomes of Random Variables $x, y, z$;

**Example 1.** Roll a die. Let $X \equiv$(the # rolled).

$P(X = 2) = 1/6$; $p(k) \equiv P(X = k) = 1/6$ for $k = 1, \ldots, 6$.

Probability mass function $p$:

| $k$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $p(k)$ | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 |

**Probability Function** $P(\cdot)$**:**

$A \to P(A)$

Properties:

- $0 \le P(A) \le 1$ for all $A$;
- $P(\Omega) = 1$;
- $P(\cup_i A_i) = \sum_i P(A_i)$ if $A_i \cap A_j = \emptyset$ for all $i \neq j$.

Call $(\Omega, \text{the events}, P)$ a probability space.

**Special case:** Symmetry. Then equally likely: $P(A) = N_A/N_\Omega$.

**Example 2.** Roll two dice. $X_1$ and $X_2$. Total $T \equiv T_2 \equiv X_1 + X_2$.

| $k_1, k_2$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 |
| 2 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 |
| 3 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 |
| 4 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 |
| 5 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 |
| 6 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 |

Here is the probability mass function $p_T(\cdot) \equiv P(T = \cdot)$:

| $k$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $36 \cdot p_T(k)$ | 1 | 2 | 3 | 4 | 5 | 6 | 5 | 4 | 3 | 2 | 1 |

Note CLT! In 7500 rolls of 2 dice, we get $T = 5$ about 833 times.

**Example 3.** Suppose that an urn contains four balls labelled 1, 2, 3, 4. Suppose that we pick a ball "at random" from the urn. Let $X$ =the # drawn, $p_k = P(X = k)$. If we draw $n$ times from the same urn, let $n_k \equiv$(the # of $k$'s in the $n$ rolls). Then we hope that $\widehat{p}_k \equiv n_k/n \to p_k$ in some sense. One of our jobs this quarter is to understand why and in what sense this happens.

**Example 4.** Suppose that an urn contains 10 balls, one ball with the label 1, two balls with the label 2, three balls with the label 3, and four balls with the label 4. Suppose we choose one ball at random from the urn. Let $X \equiv$(the # drawn). Then $X$ has the following probability mass function $p(k) \equiv P(X = k)$:

| $k$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $p(k)$ | .1 | .2 | .3 | .4 |

**Example 5.** Any $\Omega = \{a_1, \ldots, a_k\}$ and any $p_i \geq 0$ with $p_1 + \cdots + p_k = 1$ yield a probability space. Is it useful? Does it model a useful real experiment? We can study all cases at once, whether useful or not!

**Elementary Theorems:** (Kelly, pages 93-93)
- $P(A) + P(A^c) = 1$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $P(A \cup B \cup C) = \cdots$ (inclusion - exclusion principle; see page 86).
- If $B_1, \ldots, B_n$ is a (disjoint!) partition of $\Omega$, then

$A = A \cap B_1 \cup \cdots \cup A \cap B_n \equiv AB_1 + \cdots + AB_n$,
$P(A) = P(A \cap B_1) + \cdots + P(A \cap B_n) = P(AB_1) + \cdots + P(AB_n)$.

**Definition.** The conditional probability of $A$ given $B$, $P(A|B)$, is defined by

$$P(A|B) \equiv \frac{P(A \cap B)}{P(B)} \qquad \text{if } \; P(B) > 0\,.$$

**Theorem.** $P(AB) = P(A|B)P(B)$ always.

**Definition.** $A$ is "independent" of $B$ if $P(A|B) = P(A)$.

**Theorem.** $P(AB) = P(A)P(B)$ is $A$ and $B$ are independent.

Why does this definition make sense? Because, in the equally likely case

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{N_{AB}/N_\Omega}{N_B/N_\Omega} = \frac{N_{AB}}{N_B}$$

which makes sense.

**Example 1.** Draw two cards at random from a deck of 52 cards. Then

$$P(\text{first two cards are diamonds}) = P(D_1 D_2) = P(D_1)P(D_2|D_1) = \frac{13}{52}\frac{12}{51} = .0588\,.$$

Note: we solved two problems on trivial sample spaces, instead of one harder problem on a more complicated probability space.

**Fundamental Fact:** Conditional probability allows us to **revisualize** the problem.

**Example 2.** Suppose that an urn contains 6 white and 9 black balls. Suppose we draw four balls without replacement from the urn. Then

$$P(WWBB) = \frac{6}{15}\frac{5}{14}\frac{9}{13}\frac{8}{12} \approx .0659\,.$$

This is trivial with conditional probability, using 4 sample spaces. Worked on 1 sample space it is "more complicated".

**Example 3.** Rolling two dice: let $X_1 \equiv$ (# on first die), $X_2 \equiv$ (# on second die), $T \equiv X_1 + X_2$. Then

$$P(T \le 3 | T \ne 7) = \frac{3}{30} = \frac{\frac{1}{36} + \frac{2}{36}}{1 - \frac{6}{36}}\,,$$

where the first computation can be obtained by "thinking conditionally", and the second computation comes from the definition of conditional probability and direct computation.

**Example 4.** Pick one of three urns $B_1$, $B_2$, $B_3$ at random, then pick a ball at random from that urn. Let $A \equiv$ [the ball picked is R]. Suppose that urn $1 = B_1$ contains $3R, 2W$; urn $2 = B_2$ contains $7R, 3W$; urn $3 = B_3$ contains

$4R, 1W$. Then

$$\begin{aligned} P(A) &= P(AB_1) + P(AB_2) + P(AB_3) \\ &= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3) \\ &= \frac{3}{5}\frac{1}{3} + \frac{7}{10}\frac{1}{3} + \frac{4}{5}\frac{1}{3} \\ &= \frac{21}{10}\frac{1}{3} = .7\,. \end{aligned}$$

Given the event $A$, what is the probability that it was drawn from urn 1 $= B_1$? This is just

$$P(B_1|A) = \frac{P(B_1A)}{P(A)} = \frac{6/30}{21/30} = \frac{6}{21} = \frac{2}{7}\,.$$

This is an example of what is often called "Bayes formula": The probabilities $P(B_i) = 1/3$ are called the "prior probabilities", while the probabilities $P(B_i|A)$ are called the "posterior probabilities".

**Example 5.** $X_1 \equiv$(# on Red die), $X_2 \equiv$(# on White die). Then

$$P(X_1 \geq 5 \text{and } 2 \leq X_2 \leq 4) = \frac{2 \times 3}{6 \times 6} = \frac{2}{6} \cdot \frac{3}{6} = P(X_1 \geq 5)P(2 \leq X_2 \leq 4)\,.$$

## A.4. How to count; or, Combinatorics

$m \cdot n$ **Rule:** If there are $n$ ways to do step II for each of the $m$ ways to do step I, then there are $m \cdot n$ ways to do both steps.

**Sampling with and without replacement:**
$10 \cdot 10 \cdot 10 \cdots$ versus $10 \cdot 9 \cdot 8 \cdots$ .

$n! =$(# of permutations of $n$ distinct objects) $= n(n-1)\cdots 2 \cdot 1$.
$p_k^n \equiv (n)_k \equiv$ (# of $k$ long permutations of $n$ distinct objects) $= n(n-1)\cdots(n-k+1)1 = n!/(n-k)!$.
$C_k^n \equiv \binom{n}{k} \equiv$(# of subsets, or combinations, of size $k$ from $n$ distinct objects) $n!/(k!(n-k)!)$

**Example 1.** (# of batting orders using 9 starters) $= 9! = 362{,}880$.
(# of batting orders using all 26 on the roster) $= p_9^{26} = 1.3383 \times 10^{12}$.

**Example 2.** (# of poker hands) $= C_5^{52} = \binom{52}{5} = 2,598,960$
(# of bridge hands) $= C_{13}^{52} = \binom{52}{13} = 635,013,559,601$.

**Example 3.** $T$ =(# of honor cards in a bridge hand) $= X_1 + \ldots + X_13$ where $X_i = 1$ or 0 as the ith card drawn is an honor card or not. Then

$$P(T=k) = \frac{C_k^{16} C_{13-k}^{3}6}{C_{13}^{52}}, \qquad k = 0, \ldots, 13\,.$$

These probabilities are given in the following table:

| $k$ | $P(T=k)$ |
|---|---|
| 0 | .0036 |
| 1 | .0315 |
| 2 | .1135 |
| 3 | .2242 |
| 4 | .2698 |
| 5 | .2081 |
| 6 | .1053 |
| 7 | .0351 |
| 8 | .0076 |
| 9 | .0011 |
| 10 | .0001 |
| 11 | $4.3 \times 10^{-6}$ |
| 12 | $1.0 \times 10^{-7}$ |
| 13 | $8.8 \times 10^{-9}$ |

Plotting these probabilities shows that there is a "central limit effect" even after adding 13 of these dependent indicator variables, and will we see that this setting yields a CLT for dependent random variables.

**Sampling Without Replacement:** (Hypergeometric distribution.) Suppose that we are sampling *without* replacement from an urn containing $R$ red balls and $W$ white balls. Suppose that we draw $n$ times from the urn, and let $X_i$ be 1 or 0 according as the $i-$th ball drawn is Red or White. These indicator variables are *dependent*. Let

$$T \equiv T_n \equiv X_1 + \ldots + X_n = (\text{\# of Reds in the sample of size } \; n)\,.$$

Then the distribution of $T$ is the Hypergeometric$(R, N, n)$ distribution:

$$P(T=k) = \frac{\binom{R}{k}\binom{W}{n-k}}{\binom{N}{n}}$$

for each possible value of $k$; often $k = 0, \ldots, n$. Here $N \equiv R + W$ is the total number of balls in the urn.

**Example 4.** (Roll a 10-sided die twice.) Outcomes $X_1$ and $X_2$. Let $A = \{3, 4, 5\} \equiv$ "success". Then $P(X_1 \in A) = 3/10 = P(X_2 \in A) \equiv p$, and

$$P(X_1 \in A, \text{ and } X_2 \in A) = P(X_1 \in A)P(X_2 \in A) = p \cdot p = p^2 .$$

**Samping with replacement:** (Binomial distribution.) Suppose that we are sampling *with* replacement from an urn containing $R$ red balls and $W$ white balls. Suppose that we draw $n$ times from the urn, and let $X_i$ be 1 or 0 according as the $i-$th ball drawn is Red or White. These indicator variables are *independent*. Let

$$T \equiv T_n \equiv X_1 + \ldots + X_n = (\#\text{ of Reds in the sample of size } \ n) .$$

Then the distribution of $T$ is the Binomial$(n, p)$ distribution with $p = R/(R+W)$:

$$P(T = k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, \ldots, n;$$

**Definitions for Continuous Distributions**

**Definition.** A **probability density function** $f$ satisfies:
(a) $f(x) \geq 0$ for all $x \in R$;
(b) $\int_{-\infty}^{\infty} f(x)dx = 1$;
(c) for any "reasonable" set $A$, $P(A) = \int_A f(x)dx$.

**Example 1.** $X \sim$Uniform$(0, 1)$ if

$$f(x) = \left\{ \begin{array}{ll} 1, & 0 \leq x \leq 1 \\ 0, & \text{otherwise} \end{array} \right\} = 1_{[0,1]}(x) .$$

**Example 2.** $X \sim$Uniform$(a, b)$ if

$$f(x) = \left\{ \begin{array}{ll} 1/(b-a), & a \leq x \leq b \\ 0, & \text{otherwise} \end{array} \right\} = \frac{1}{b-a} 1_{[a,b]}(x) .$$

**Example 3.** $X \sim$ exponential($\lambda$) if the density function $f \equiv f_X$ for $X$ is given by

$$f(x) = \lambda \exp(-\lambda x) 1_{[0,\infty)}(x) \,.$$

**Example 4.** $X \sim$N$(0,1)$ if the density function $f$ for $X$ is given by

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \equiv \phi(x), \qquad -\infty < x < \infty \,.$$