1. Review of Probability.

What is probability? Perform an "experiment". The result is not predictable. One of finitely many possibilities $R_1, R_2, \cdots, R_k$ can occur. Some are perhaps more likely than others. We assign nonnegative numbers $p_i = P[R_i]$ such that $p_i \geq 0$ and $\sum_i p_i = 1$. The interpretation is that we know (from experience ?) that, if we repeated the experiment a large number of times, these events would occurs more or less in these these prportions. In other words if we repeat the experiment N times, for large N,

$$\frac{\#(R_i)}{N} \simeq p_i$$

Often when there is no reason to prefer one over the others we may set $P(R_i) = \frac{1}{k}$.

1. Examples. Toss a coin. H or T. $P(H) = P(T) = \frac{1}{2}$

2. Throw a die. $1, 2, \ldots, 6$. $P(1) = P(2) = \cdots = P(6) = \frac{1}{6}$

Repeated Experiments.

Toss a coin twice. $P(HH) = P(HT) = P(TH) = P(TT) = \frac{1}{4}$.

Independence.

If $P[R_i] = p_i$, then if we repeat twice under the assumption of independence we have $P[R_i R_j] = P[R_i]P[R_j]$.

They can be different experiments. $P[R_i S_j] = P[R_i]P[S_j]$

You can have many experiments that are mutually independent.

For example for any string of length n, $P[HHTTHTTH \cdots TT] = \frac{1}{2^n}$

Absractly $\mathcal{X}$ is a finite set of points $\{x\}$ or $\{x_1, \ldots, x_k\}$ and $\{p(x)\}$ or $\{p_i\}$ are numbers adding up to 1. We extend the definition to subsets $A \subset X$.

$$P(A) = \sum_{i: x_i \in A} p_i = \sum_{i: x \in A} p(x)$$

P has the properties $0 \leq P(A) \leq 1$. $P(X) = 1$. $P(\emptyset) = 0$ and if $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$.

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

Mappings. $F : X \to Y$. There is a natural Q on Y defined by

$$Q(B) = P(F^{-1}A) = \sum_{i:F(x_i)\in B} p_i = \sum_{x:F(x)\subset B} p(x)$$

If $Y = \mathbf{R}$ then F is called random variable. Example $x = \{HHTT \cdots H\}$ a string of length n. $p(x) = \frac{1}{2^n}$. F is the number of heads. $\{x : F(x) = r\}$ has $\binom{n}{r}$ strings in it. So

$$P[F = r] = \frac{\binom{n}{r}}{2^n}$$

The heads and tails may have unequal probabilities. $P(H) = p$ and $P(T) = 1 - p$. Then

$$p(x) = p^{F(x)}(1-p)^{n-F(x)}$$

Therefore

$$P[F = r] = \binom{n}{r} p^r (1-p)^{n-r}$$

Expectations. If $X \subset \mathbf{R}$, then the mean of the distribution p is defined as

$$m = \sum_x xp\,(x)$$

More generally if F is random variable then

$$E[F(x)] = \sum_x F(x)p(x) = \sum_y yq(y)$$

where $q(y) = P[F(x) = y] = \sum_{x:F(x)=y} p(x)$.

If F and G are random variables then

$$E[aF + bG] = aE[F] + bE[G]$$

If P is on X and Q is on Y, then on $Z = X \times Y$ $R = P \times Q$ is the product distribution defined by $r(\{x, y\}) = p(x)q(y)$. Then it is easy to verify that

$$\begin{aligned} E[F(x)G(y)] &= \sum_{x,y} F(x)G(y)p(x)q(y) = [\sum_x F(x)p(x)] \times [\sum_y G(y)q(y)] \\ &= E[F(x)]E[G(y)] \end{aligned}$$

For the Binomial Distribution

$$m = \sum_r r \binom{n}{r} p^r (1-p)^{n-r} = np$$

$F = F_1 + F_2 + \cdots + F_n$. Each $F_i = 1$ or 0 with probability p and $E[F_i] = p$ and $E[F] = n\,p$.

Waiting Times; Suppose we have independent tosses with $P(H) = p$ and $P(T) = 1 - p$, F is the number of tries before a Head shows up, (including the last one), then we obtain the Geometric distribution.

$$P[F = r] = p(1-p)^{r-1}$$

For the Geometric Distribution

$$E[F] = \sum pr(1-p)^{r-1} = \frac{1}{p}$$

Could you have guessed it?

Conditional Probability.

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

Example. Drawing without replacement. We have an urn containing r red and g green ball. A ball is drawn at random. Then another ball is drawn at random with out replacement. A is the event that the first ball is red. B is the event that the second ball is green. It is clear that

$$P(B|A) = \frac{g}{g + r - 1}$$

What about $P(A|B)$?

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B|A)}{P(B)} = \frac{r}{r + g - 1}$$

Bayes' rule; B_i are disjoint and there union is X the whole space. Then

$$P(A) = \sum_i P(B_i \cap A) = \sum_i P(B_i)P(A|B_i)$$

In the previous example

$$\begin{aligned} P(B) &= P(B|A)P(A) + P(B|A^c)P(A^c) \\ &= \frac{g}{g+r-1}\frac{r}{g+r} + \frac{g-1}{g+r-1}\frac{g}{g+r} \\ &= \frac{g}{g+r} \end{aligned}$$

Conditional Expectation.

If $F(x)$ is a random variable on $\mathcal{X} = \{x\}$ with probbailities $\{p(x)\}$, the expectation of F on any A can be defined as

$$E[F|A] = \frac{\sum_{x\in A} F(x)p(x)}{\sum_{x\in A} p(x)} = \frac{1}{P(A)} \sum_{x\in A} F(x)p(x)$$

If X and Y are two random variables then

$$E[Y|X] = f(X)$$

where $f(x)$ is defined for every x with $P[X = x] > 0$ by the formula

$$\begin{aligned} f(x) &= E[Y|X = x] \\ &= \sum yP[Y = y|X = x] \\ &= \frac{1}{P[X = x]} \sum_y yP[X = x, Y = y] \end{aligned}$$

It is easy to check that

$$E[Y] = E[f(X)] = E[E[Y|X]]$$

Mean and Variance.

Let $\mathcal{X} = \{x\}$ be a be a finite set with associated probabilities $\{p(x)\}$. We saw that if $Y = f(x)$ is a random variable then, with $q(y) = P[Y = y] = \sum_{x:y=f(x)} p(x)$,

$$E[Y] = E[f(x)] = \sum yq(y) = \sum_x f(x)p(x)$$

We can similarly define $E[Y^2] = \sum_y y^2 q(y)$. The Variance of Y is defined as

$$E[[Y - E[Y]]^2] = E[Y^2] - 2[E[Y]]^2 + [E[Y]]^2 = E[Y^2] - [E[Y]]^2$$

Measures the spread. $V(aX + b) = a^2 V(X)$.

If X and Y are independent random variables the $Var(X+Y) = Var(X) + Var(Y)$. If we expand $[X - E(X) + Y - E(Y)]^2$ we get an additional cross term $2[X - E(X)][Y - E(Y)]$ and if X and Y are independent

$$E[[X - E(X)][Y - E(Y)]] = E[[X - E(X)]] \times E[[Y - E(Y)]] = 0$$

Some important Discrete distributions.

1. Binomial Distribution . $\{1, 2, \ldots, n\}$. $P(r) = \binom{n}{r} p^r (1-p)^{n-r}$. Mean $= np$. Variance $= np(1-p)$.

One way to compute is the use of generating functions.

$$E[e^{\theta X}] = M(\theta) = \sum_{r=0}^{\infty} \frac{\theta^r E[X^r]}{r!} = \sum_r e^{\theta r} \binom{n}{r} p^r (1-p)^{n-r} = (pe^\theta + q)^n$$

$$M'(\theta) = n(pe^\theta + 1 - p)^{n-1} pe^\theta$$

$$M''(\theta) = n(n-1)(pe^\theta + 1 - p)^{n-2} p^2 e^{2\theta} + n(pe^\theta + 1 - p)^{n-1} pe^\theta$$

Mean $= M'(0) = np$.

Variance $= M''(0) - [M'(0)]^2 = n(n-1)p^2 + np - n^2p^2 = np - np^2 = np(1-p)$

2. Geometric Distribution. $\{0, 1, 2, \ldots, ...\}$. $P(r) = p(1-p)^{r-1}$

$$M(\theta) = \sum_{r \geq 1} p(1-p)^{r-1} e^{r\theta} = \frac{pe^{\theta}}{1-(1-p)e^{\theta}} = \frac{p}{p + e^{-\theta} - 1}$$

$$M'(0) = \frac{1}{p}$$

$$M''(0) = \frac{2}{p^2} - \frac{1}{p}$$

Mean $= \frac{1}{p}$. Variance $= \frac{1}{p^2} - \frac{1}{p} = \frac{1-p}{p^2}$.

2. Poisson Distribution. $\{0, 1, 2, \ldots, ...\}$. $P(r) = e^{-\lambda} \frac{\lambda^r}{r!}$

$$M(\theta) = e^{-\lambda} e^{\lambda e^{\theta}}$$

Mean $= M'(0) = \lambda$, Variance $= \lambda$. Binomial $p << 1$, $n >> 1$ $np = \lambda$, then as $n \to \infty$,$p \to 0$, $np \to \lambda$

$$\binom{n}{r} p^r (1-p)^{n-r} \to e^{-\lambda} \frac{\lambda^r}{r!}$$

$np(1-p) \to \lambda$.

Sums of Independent Random variables.

$P[X = r] = p(r)$ $P[Y = r] = q(r)$. X and Y are independent.

$\pi(r) = P[X + Y = r] = \sum_{a+b=r} p(a)q(b)$. $\pi = p * q$ is the convolution of p and q.

Probability generating functions. $P(z) = \sum p(r) z^r$. Replace e^{θ} by z.

Binomial: $(p + qz)^n$

Geomteric: $\frac{pz}{1-(1-p)z}$

Poisson: $e^{\lambda(z-1)}$.

$Bin(n, p) * Bin(m, p) = Bin(n + m, p)$

$Poisson(\lambda) * Poisson(\mu) = Poisson(\lambda + \mu)$

Negative Binomial: Convolutions of Geometric. $\left[\frac{pz}{1-(1-p)z}\right]^n$

$P_n[X = n + r] = \binom{n+r-1}{r}(1-p)^r p^n$

Distribution functions. $F_X(t) = P[X \leq t] = \sum_{x \leq t} P[X = x]$

If X and Y are independent and $Z = \max\{X, Y\}$, then $F_Z(t) = F_X(t)F_Y(t)$.

Assignment 1.

k dice are thrown. Assume that all the sides have the same probability of showing up and the scores $\{X_1, \ldots, X_k\}$ of the k dice are independent. What is the probability distribution of $F = \max_{1 \le i \le k} X_i$? Calculate $E[F]$ and $V(F)$. What happens when k is large?