# AP STATISTICS
# TOPIC V: RANDOM VARIABLES

PAUL L. BAILEY

Within this document, we will assume that all probability spaces are finite.

## 1. Random Variables

**Definition 1.** Let $S$ be probability space. A *random variable* on $S$ is a function

$$X : S \to \mathbb{R}.$$

A random variable on a probability space $S$ induces the structure of a probability space on the image, as follows. Let $S$ be a probability space, $X : S \to \mathbb{R}$ a random variable, and $I = \text{range}(X)$. Note that if $S$ is finite, then so is $I$. For each point $x \in I$, assign the probability $f_X(x)$ to be the probability of the preimage of $x$ under $X$.

**Definition 2.** Let $X : S \to \mathbb{R}$ be a random variable. The *probability density function* (pdf) of $X$ is

$$f_X : \mathbb{R} \to [0,1] \quad \text{given by} \quad f_X(x) = P(X^{-1}(x)).$$

This function is also known as the *probability mass function* (pmf).

Let $X : S \to \mathbb{R}$ be a random variable. The *cumulative density function* (cdf) of $X$ is

$$F_X : \mathbb{R} \to [0,1] \quad \text{given by} \quad F_X(x) = P(X^{-1}((-\infty, x]).$$

Although the notation $f_X$ is standard, we will more frequently use the following notation, which is also standard.

- $P(X = x) = P(X^{-1}(x))$
- $P(X \leq x) = P(X^{-1}((-\infty, x])$
- $P(x_1 \leq X \leq x) = P(X^{-1}([x_1, x_2])$

**Proposition 1. (Dirty Trick Theorem)**
*Let $X : S \to \mathbb{R}$ be a random variable. Then*

$$\sum_{x \in \mathbb{R}} P(X = x) = 1.$$

**Definition 3.** Let $X$ and $Y$ be random variables on $S$. We say that $X$ and $Y$ are *independent* if, for every $x, y \in \mathbb{R}$,

$$P(\{s \in X \mid X(s) = x \text{ and } Y(s) = y\}) = P(X = x) \cdot P(Y = y).$$

---

*Date*: September 19, 2023.

## 2. Expectation

**Definition 4.** Let $X : S \to \mathbb{R}$ be a random variable. The *expectation* of $X$ is a real number
$$E(X) = \sum_{x \in \mathbb{R}} xP(X = x).$$

**Proposition 2.** *Let $S$ be a finite uniform probability space, and let $X : S \to \mathbb{R}$ be a random variable. Then*
$$E(X) = \frac{1}{|S|} \sum_{s \in S} X(s).$$

*Proof.* We view the $X$ as producing a statistical variable on the population $S$, with mean $\mu$. Let $E_x = X^{-1}(x)$ denote the event the $X = x$; then $|E_x|$ is the number of members of $S$ which map to $x$, and we have
$$\begin{aligned}
\mu &= \frac{1}{|S|} \sum_{s \in S} X(s) \\
&= \frac{1}{|S|} \sum_{x \in \mathbb{R}} x|E_x| \\
&= \sum_{x \in \mathbb{R}} x \frac{|E_x|}{|S|} \\
&= \sum_{x \in \mathbb{R}} xP(X = x) \\
&= E(x).
\end{aligned}$$

□

That is, the expectation of a random variable on a finite uniform probability space is the average value of the random variable. Thus if we let $\mu = E(X)$, we arrive at the mean of the population's values.

**Proposition 3** (Linearity of Expectation)**.** *Let $X$ and $Y$ be random variables, and let $a \in \mathbb{R}$. Then*

- **(a)** $E(X + Y) = E(X) + E(Y)$*;*
- **(b)** $E(aX) = aE(X)$.

**Proposition 4** (Independence of Expectation)**.** *Let $X$ and $Y$ be independent random variables. Then*
$$E(XY) = E(X)E(Y).$$

## 3. Variance

**Definition 5.** Let $S$ be a finite probability space and let $X : S \to \mathbb{R}$ be a random variable on $S$. Let $\mu = E(X)$. The *variance* of $X$ is

$$V(X) = \sum_{x \in \mathbb{R}} (x - \mu)^2 P(X = x).$$

**Proposition 5.** *Let $S$ be a finite probability space and let $X : S \to \mathbb{R}$ be a random variable on $S$. Then*

$$V(X) = E(X^2) - (E(X))^2.$$

*Proof.* Let $\mu = E(X)$. Then

$$\begin{aligned} V(X) &= \sum_{x \in \mathbb{R}} (x - \mu)^2 P(X = x) \\ &= \sum_{x \in \mathbb{R}} x^2 P(X = x) - 2\mu \sum_{x \in \mathbb{R}} x P(X = x) + \mu^2 P(X = x) \\ &= \sum_{x \in \mathbb{R}} x^2 P(X = x) - 2\mu \sum_{x \in \mathbb{R}} x P(X = x) + \mu^2 \\ &= \sum_{x \in \mathbb{R}} x^2 P(X = x) - 2\mu^2 + \mu^2 \\ &= \sum_{x \in \mathbb{R}} x^2 P(X = x) - \mu^2 \\ &= E(X^2) - (E(X))^2. \end{aligned}$$

□

We recall that the variance of a variable is $\sigma^2 = \dfrac{\sum (x - \mu)^2}{N}$, where $N$ is the size of the population. If we apply this in our current context,

$$\sigma^2 = \frac{\sum_{s \in S} (X(s) - \mu)^2}{|S|} = \sum_{x \in \mathbb{R}} (x - \mu)^2 P(X = x).$$

Thus we set $\mu(X) = E(X)$ and $\sigma(X) = \sqrt{V(X)}$.

**Proposition 6.** *Let $S$ be a finite probability space. Let $X$ and $Y$ be independent random variables on $S$ and let $a, b \in \mathbb{R}$. Then*

$$V(aX + bY) = a^2 V(X) + b^2 V(Y).$$

*Proof.*

$$\begin{aligned} V(aX + bY) &= E((aX + bY)^2) - E(aX + bY)^2 \\ &= E(a^2 X^2 + 2abXY + b^2 Y^2) - (aE(X) + bE(Y))^2 \\ &= a^2 E(X^2) + 2abE(XY) + b^2 E(Y^2)) - (a^2 E(X)) \end{aligned}$$

□

## 4. Seven Great Discrete Distributions

We now describe the seven great discrete distributions:

**(1)** Uniform Distribution
**(2)** Binomial Distribution
**(3)** Geometric Distribution
**(4)** Poisson Distribution
**(5)** Hypergeometric Distribution (Not on AP Exam)
**(6)** Wilcoxon Distribution (Not on AP Exam)
**(7)** Survey Distribution

**Great Discrete Distribution 1. Uniform Distribution**
Let $S$ be a finite set of cardinality $n$, and form the uniform probability space $(S, \mathcal{P}(S), P)$, where $P : \mathcal{P}(S) \to [0, 1]$ is given by $P(E) = \frac{|E|}{|S|} = \frac{|E|}{N}$.

Let $X : S \to \{1, \dots, N\}$ be a bijective function. Then $X$ is a discrete random variable. We say that $X$ has a *uniform distribution.*

The image of $X$ is $\{1, \dots, N\}$.

The density of $X$ is

$$P(X = x) = \begin{cases} \frac{1}{N} & \text{if} \quad x = \text{img}(X); \\ 0 & \text{otherwise .} \end{cases}$$

The expectation of $X$ is

$$E(X) = \frac{N+1}{2}.$$

*Proof.* Thus

$$\begin{aligned} E(X) &= \sum_{x \in \mathbb{R}} x P(X = x) && \text{definition of expectation} \\ &= \sum_{k=1}^{N} k \cdot \frac{1}{N} && \text{definition of uniform distribution} \\ &= \frac{1}{N} \sum_{k=1}^{N} k && \text{since } N \text{ is constant with respect to } k \\ &= \frac{1}{N} \left( \frac{N(N+1)}{2} \right) && \text{sum of an arithmetic series} \\ &= \frac{N+1}{2}. \end{aligned}$$

□

**Great Discrete Distribution 2. Binomial Distribution**
Let $S$ be a finite set of cardinality $N$, and form the uniform probability space $(S, \mathcal{P}(S), P)$, where $P : \mathcal{P}(S) \to [0, 1]$ is given by $P(E) = \frac{|E|}{|S|} = \frac{|E|}{N}$.

Let $R \subset S$ with $|R| = r$ and let $p = P(R) = \frac{r}{N}$.

Define a discrete random variable $Y : S \to \mathbb{R}$ by

$$Y(s) = \begin{cases} 1 & \text{if } s \in R; \\ 0 & \text{if } s \notin R. \end{cases}$$

We say that $Y$ is the *bernoulli* random variable associated to the event $R$.

The density of $Y$ is

$$P(Y=y)=\begin{cases} p & \text{if } y=1; \\ 1-p & \text{if } y=0; \\ 0 & \text{otherwise.} \end{cases}$$

Let $n$ be a positive integer. Let $T=\times_{i=1}^{n} S$, the cartesian product of $S$ with itself $n$ times. Then $|T|=N^n$. Form the uniform probability space $(T, \mathcal{P}(T), Q)$, where for $F\subset T$ we have $Q(F)=\frac{|Q|}{|T|}=\frac{|F|}{N^n}$.

Define a discrete random variable $X: T\to\mathbb{R}$ by

$$X(s_1,\dots,s_n)=\sum_{i=1}^{n} Y(s_i).$$

We say that $X$ has a *binomial distribution.*

The image of $X$ is

$$\operatorname{img}(X)=\{0,1,2,\dots,n\}.$$

The density of $X$ is

$$P(X=x)=\binom{n}{x}p^x(1-p)^{n-x}.$$

The expectation of $X$ is

$$E(X)=np.$$

*Proof.* Let $q=1-p$. Note that $(n-1)-(k-1)=n-k$, so

$$k\binom{n}{k} \quad = \quad k\frac{n!}{k!(n-k)!} \quad = \quad n\frac{(n-1)!}{(k-1)!(n-k)!} \quad = \quad n\binom{n-1}{k-1},$$

Thus

$$\begin{aligned}
E(X) &= \sum_{x\in\mathbb{R}} xP(X=x) && \text{definition of expectation} \\
&= \sum_{k=0}^{n} k\binom{n}{k}p^k q^{n-k} && \text{definition of binomial distribution} \\
&= \sum_{k=1}^{n} k\binom{n}{k}p^k q^{n-k} && \text{since for } k=0,\ k\binom{n}{k}p^k q^{n-k}=0 \\
&= \sum_{k=1}^{n} n\binom{n-1}{k-1}p^k q^{n-k} && \text{since } k\binom{n}{k}=n\binom{n-1}{k-1} \\
&= np\sum_{k=1}^{n} \binom{n-1}{k-1}p^{k-1} q^{n-k} && \text{factor out } np \\
&= np\sum_{j=0}^{m} \binom{m}{j}p^{j} q^{m-j} && \text{put } m=n-1 \text{ and } j=k-1 \\
&= np(p+q)^n && \text{Binomial Theorem} \\
&= np. && \text{since } p+q=1
\end{aligned}$$

□

The variance of $X$ is

$$V(X) = npq.$$

*Proof.* We know that $V(X) = E(X^2) - (E(X))^2$. By definition, $E(X^2) = \sum_{x\in\mathbb{R}} x^2 P(X = x)$. Let $q = 1 - p$, so that $p + q = 1$. Then

$$\begin{aligned}
E(X^2) &= \sum_{k=0}^{n} k^2 \binom{n}{k} p^k q^{n-k} \\
&= \sum_{k=0}^{n} kn \binom{n-1}{k-1} p^k q^{n-k} \\
&= np \sum_{k=1}^{n} k \binom{n-1}{k-1} p^{k-1} q^{(n-1)-(k-1)} \\
&= np \sum_{j=0}^{m} (j+1) \binom{m}{j} p^j q^{m-j} \qquad \text{where } m = n-1 \text{ and } j = k-1 \\
&= np \left( \sum_{j=0}^{m} j \binom{m}{j} p^j q^{m-j} + \sum_{j=0}^{m} \binom{m}{j} p^j q^{m-j} \right) \\
&= np \left( \sum_{j=0}^{m} m \binom{m-1}{j-1} p^j q^{m-j} + \sum_{j=0}^{m} \binom{m}{j} p^j q^{m-j} \right) \\
&= np \left( (n-1)p \sum_{j=0}^{m} \binom{m-1}{j-1} p^{j-1} q^{(m-1)-(j-1)} + \sum_{j=0}^{m} \binom{m}{j} p^j q^{m-j} \right) \\
&= np \left( (n-1)p(p+q)^{m-1} + (p+q)^m \right) \\
&= np((n-1)p + 1) \\
&= n^2p^2 + np(1-p) \\
&= npq + n^2p^2
\end{aligned}$$

Thus

$$\begin{aligned}
V(X) &= E(X^2) - (E(X))^2 \\
&= npq + n^2p^2 - (np)^2 \\
&= npq.
\end{aligned}$$

□

**Great Discrete Distribution 3. Geometric Distribution**
Let $S$ be a finite set of cardinality $N$, and form the uniform probability space $(S, \mathcal{P}(S), P)$, where $P : \mathcal{P}(S) \to [0, 1]$ is given by $P(E) = \frac{|E|}{|S|} = \frac{|E|}{N}$.

Let $R \subset S$ with $|R| = r$ and let $p = P(R) = \frac{r}{N}$. Let $Y : S \to \mathbb{R}$ be the bernoulli random variable associated to $R$, so that

$$Y(s) = \begin{cases} 1 & \text{if } s \in R; \\ 0 & \text{if } s \notin R. \end{cases}$$

Let $T$ be the set of all sequences in $S$, so that

$$T = \{\sigma : \mathbb{N} \to S\}.$$

We wish to put a probability measure on $T$; however, $T$ is an uncountable set. Let $\mathcal{E}$ be the sigma algebra generated by the sets

$$E_n(\tau) = \{\sigma \in T \mid \sigma(i) = \tau(i) \text{ for all } i > n\}.$$

Define $Q(E_n(\tau)) = \frac{1}{N^n}$.

Define a discrete random variable $X : T \to \mathbb{R}$ by

$$X(\sigma) = \begin{cases} \min\{i \in \mathbb{N} \mid Y(\sigma(i)) = 1\} & \text{if this set is nonempty;} \\ 0 \quad \text{otherwise.} \end{cases}$$

We say that $X$ has a *geometric* distribution.

The range of $X$ is

$$\operatorname{img}(X) = \{0, 1, 2, \dots\}.$$

The density of $X$ is

$$f_X(x) = \begin{cases} p(1-p)^{x-1} & \text{if} \quad x \in \{1, 2, \dots\}; \\ 0 \quad \text{otherwise.} \end{cases}$$

The expectation of $X$ is

$$E(X) = \frac{1}{p}.$$

*Proof.* We have

$$\begin{aligned}
E(X) &= \sum_{x\in\mathbb{R}} xP(X = x) \\
&= \sum_{k=1}^{\infty} kP(X = k) \\
&= \sum_{k=0}^{\infty} kp(1-p)^{k-1} \\
&= p\sum_{k=0}^{\infty} k(1-p)^{k-1} \\
&= p\sum_{k=0}^{\infty} (-\frac{d}{dp}(1-p)^{k}) \\
&= -p \cdot \frac{d}{dp}\sum_{k=0}^{\infty}(1-p)^{k} \\
&= -p \cdot \frac{d}{dp}\frac{1}{1-(1-p)} \\
&= -p \cdot \frac{d}{dp}\frac{1}{p} \\
&= -p \cdot \frac{-1}{p^2} \\
&= \frac{1}{p}.
\end{aligned}$$

□

The variance of $X$ is

$$V(X) = \frac{1-p}{p^2}.$$

*Proof.* We have

$$\begin{aligned}
E(X^2) &= \sum_{x\in\mathbb{R}} x^2 P(X = x) \\
&= \sum_{k=1}^{\infty} k^2 P(X = k) \\
&= \sum_{k=0}^{\infty} k^2 p(1-p)^{k-1} \\
&= p\sum_{k=0}^{\infty} k(k+1)(1-p)^{k-1} - \sum_{k=0}^{\infty} kp(1-p)^{k-1} \\
&= p\sum_{k=0}^{\infty}(\frac{d^2}{dp^2}(1-p)^{k+1}) - E(X) \\
&= p\cdot\frac{d^2}{dp^2}\sum_{k=1}^{\infty}(1-p)^k - \frac{1}{p} \\
&= p\cdot\frac{d^2}{dp^2}(\frac{1}{1-(1-p)} - 1) - \frac{1}{p} \\
&= p\cdot\frac{d^2}{dp^2}(\frac{1}{p} - 1) - \frac{1}{p} \\
&= p\cdot\frac{2}{p^3} - \frac{1}{p} \\
&= \frac{2-p}{p^2}.
\end{aligned}$$

Thus

$$V(X) = E(X^2) - (E(X))^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$

□

**Okay Discrete Distribution 3. Truncated Geometric Distribution**
Let $S$, $R$, and $Y$ be as above. Let $T$ be the cartesian product of $S$ with itself $n$ times. Define a discrete random variable $X : T \to \mathbb{R}$ by

$$X(s_1, \ldots, s_n) = \begin{cases} \min\{i \leq n \mid Y(s_i) = 1\} & \text{if this set is nonempty;} \\ 0 & \text{otherwise.} \end{cases}$$

We say that $X$ has a *truncated geometric* distribution.

**Great Discrete Distribution 4. Poisson Distribution**
Let $T$ be an infinite probability space and let $X : T \to \mathbb{R}$ be a random variable whose density function satisfying the following.

The image of $X$ is

$$\operatorname{img}(X) = \{0, 1, 2, 3, \dots\}.$$

The density of $X$ is

$$f_X(x) = \begin{cases} e^{-\lambda} \frac{\lambda^x}{x!} & \text{for } x \in \operatorname{img}(X); \\ 0 & \text{otherwise.} \end{cases}$$

We say that $X$ has a *Poisson distribution*.

The expectation of $X$ is

$$E(X) = \lambda.$$

*Proof.* Consider that

$$\begin{aligned} E(X) &= \sum_{x \in \mathbb{R}} x P(X = x) \\ &= \sum_{k=0}^{\infty} k P(X = k) \\ &= \sum_{k=1}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} \\ &= \frac{\lambda}{e^{\lambda}} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\ &= \frac{\lambda}{e^{\lambda}} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \\ &= \frac{\lambda}{e^{\lambda}} e^{\lambda} \qquad \text{using the Taylor series for } e^x \\ &= \lambda \end{aligned}$$

□

The variance of $X$ is

$$V(X) = \lambda.$$

*Proof.* Consider that

$$\begin{aligned}
E(X^2) &= \sum_{x\in\mathbb{R}} x^2 P(X=x) \\
&= \sum_{k=0}^{\infty} k^2 P(X=k) \\
&= \sum_{k=1}^{\infty} k^2 e^{-\lambda}\frac{\lambda^k}{k!} \\
&= \frac{\lambda}{e^\lambda}\sum_{k=1}^{\infty} k\frac{\lambda^{k-1}}{(k-1)!} \\
&= \frac{\lambda}{e^\lambda}\left(\sum_{k=1}^{\infty}(k-1)\frac{\lambda^{k-1}}{(k-1)!} + \sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}\right) \\
&= \frac{\lambda}{e^\lambda}\left(\lambda\sum_{k=2}^{\infty}(k-2)\frac{\lambda^{k-2}}{(k-2)!} + \sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}\right) \\
&= \frac{\lambda}{e^\lambda}\left(\lambda\sum_{i=0}^{\infty} i\frac{\lambda^i}{i!} + \sum_{j=0}^{\infty}\frac{\lambda^j}{j!}\right) \\
&= \frac{\lambda}{e^\lambda}\left(\lambda e^\lambda + e^\lambda\right) \\
&= \lambda(\lambda+1) \\
&= \lambda^2 + \lambda.
\end{aligned}$$

Thus

$$V(X) = E(X^2) - (E(X))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

□

The Poisson distribution is the limit of the binomial distribution in the following sense.

Let $p \in (0,1)$ and let $X_n$ be a random variable with binomial $(n,p)$ distribution. Then $\mu = E(n) = np$, so $p = \mu/n$. Let $\rho_n : \mathbb{R} \to \mathbb{R}$ denote the density of the $n^{\text{th}}$ binomial distribution. For $x = 0, 1, \ldots, n$, we have

$$\begin{aligned}
\rho(x) &= \binom{n}{x} p^x (1-p)^{n-x} \\
&= \frac{n(n-1)(n-2)\cdots(n-x+1)}{x!}\left(\frac{\mu}{n}\right)^x\left(1-\frac{\mu}{n}\right)^{n-x} \\
&= \frac{n}{n}\cdot\frac{n-1}{n}\cdot\cdots\cdot\frac{n-x+1}{n}\cdot\frac{\mu^x}{x!}\cdot\left(1-\frac{\mu}{n}\right)^n\left(1-\frac{\mu}{n}\right)^{-x}
\end{aligned}$$

Taking the limit as $n \to \infty$ yields

$$\rho(x) = \frac{\mu^x e^{-\mu}}{x!}.$$

It is simply traditional to use $\lambda$ as opposed to $\mu$ for the Poisson distribution.

**Great Discrete Distribution 5. Hypergeometric Distribution**
Let $S$ be a finite set of cardinality $N$, and form the uniform probability space $(S, \mathcal{P}(S), P)$, where $P : \mathcal{P}(S) \to [0, 1]$ is given by $P(E) = \frac{|E|}{N}$.

Let $R \subset S$ with $|R| = r$ and let $p = P(R) = \frac{r}{N}$. Let $Y : S \to \mathbb{R}$ be the bernoulli random variable associated to $R$, so that

$$Y(s) = \begin{cases} 1 & \text{if } s \in R; \\ 0 & \text{if } s \notin R. \end{cases}$$

The expectation of $Y$ is

$$E(Y) = p.$$

Let $n$ be an integer such that $0 \leq n \leq N$. Set

$$T = \{A \in \mathcal{P}(S) \mid |A| = n\}.$$

Then $|T| = \binom{N}{n}$. Form the uniform probability space $(T, \mathcal{P}(T), Q)$, where for $F \subset T$ we have $Q(F) = \frac{|F|}{|T|} = \frac{|F|}{\binom{N}{n}}$.

Define a random variable $X : T \to \mathbb{R}$ by

$$X(A) = \sum_{a \in A} Y(a).$$

Then $X(A) = |A \cap R|$.

The image of $X$ is

$$\operatorname{img}(X) = \{0, 1, \ldots, n\}.$$

The density of $X$ is

$$f_X(x) = \begin{cases} \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} & \text{if} \quad x \in \operatorname{img}(X); \\ 0 & \text{otherwise.} \end{cases}$$

The expectation of $X$ is

$$E(X) = \frac{nr}{N} = np.$$

Obtain this as follows. For $a \in S$, the number of sets in $T$ containing $a$ is $\binom{N-1}{n-1}$. Thus

$$\begin{aligned} E(X) &= \frac{1}{|T|} \sum_{A \in T} X(A) \\ &= \frac{1}{|T|} \sum_{A \in T} \sum_{a \in A} Y(a) \\ &= \frac{1}{|T|} \sum_{a \in R} |\{A \in T \mid a \in A\}| \\ &= \frac{1}{|T|} \sum_{a \in R} \binom{N-1}{n-1} \\ &= \frac{\binom{N-1}{n-1} r}{\binom{N}{n}} \\ &= \frac{nr}{N}. \end{aligned}$$

**Great Discrete Distribution 6. Wilcoxon Distribution**
Let $S$ be a finite set of cardinality $N$, and form the uniform probability space $(S, \mathcal{P}(S), P)$, where $P : \mathcal{P}(S) \to [0, 1]$ is given by $P(E) = \frac{|E|}{N}$.

Let $Y : S \to \{1, 2, \ldots, N\}$ be a bijective random variable.

The expectation of $Y$ is

$$E(Y) = \frac{1}{N} \sum_{i=1}^{N} i = \frac{1}{N} \cdot \frac{N(N+1)}{2} = \frac{N+1}{2}.$$

Let $n$ be an integer such that $0 \leq n \leq N$. Set

$$T = \{A \in \mathcal{P}(S) \mid |A| = n\}.$$

Then $|T| = \binom{N}{n}$. Form the uniform probability space $(T, \mathcal{P}(T), Q)$, where for $F \subset T$ we have $Q(F) = \frac{|F|}{\binom{N}{n}}$.

Define a random variable $X : T \to \mathbb{R}$ by

$$X(A) = \sum_{a \in A} Y(a).$$

We say that $X$ has a *Wilcoxon distribution*.

The image of $X$ is

$$\operatorname{img}(X) = \{\frac{n(n+1)}{2}, \frac{n(n+1)}{2} + 1, \ldots, \frac{N(N+1)}{2} - \frac{(N-n)(N-n+1)}{2}\}.$$

The density of $X$ is difficult to describe.

The expectation of $X$ is

$$E(X) = \frac{n(N+1)}{2}.$$

**Great Discrete Distribution 7. Sample Survey Distribution**
Let $S$ be a finite set of cardinality $N$, and form the uniform probability space $(S, \mathcal{P}(S), P)$, where $P : \mathcal{P}(S) \to [0, 1]$ is given by $P(E) = \frac{|E|}{N}$.

Let $Y : S \to \mathbb{R}$ be a discrete random variable.

Let $n$ be an integer such that $0 \leq n \leq N$. Set

$$T = \{A \in \mathcal{P}(S) \mid |A| = n\}.$$

Then $|T| = \binom{N}{n}$. Form the uniform probability space $(T, \mathcal{P}(T), Q)$, where for $F \subset T$ we have $Q(F) = \frac{|F|}{\binom{N}{n}}$.

Define a random variable $X : T \to \mathbb{R}$ by

$$X(A) = \sum_{a \in A} Y(a).$$

We say that $X$ has a *sample survey* distribution.

The image of $X$ is determined by the image of $Y$.

The density of $X$ is difficult to describe.

The expectation of $X$ is

$$E(X) = nE(Y).$$

Obtain this as follows.

$$\begin{aligned}
E(X) &= \frac{1}{|T|} \sum_{A \in T} X(A) \\
&= \frac{1}{|T|} \sum_{A \in T} \sum_{a \in A} Y(a) \\
&= \frac{1}{|T|} \sum_{a \in S} |\{A \in T \mid a \in A\}| \cdot Y(a) \\
&= \frac{1}{|T|} \sum_{a \in S} \binom{N-1}{n-1} Y(a) \\
&= \frac{\binom{N-1}{n-1}}{\binom{N}{n}} \sum_{a \in S} Y(a) \\
&= \frac{n}{N} \sum_{a \in S} Y(a) \\
&= nE(Y).
\end{aligned}$$

## 5. Random Vectors

**Definition 6.** Let $(S, \mathcal{E}, P)$ be a probability space. A function $\vec{X} : S \to \mathbb{R}^n$ is called a *random vector* if $\vec{X}^{-1}((-\infty, a]^n) \in \mathcal{E}$ for every $a \in \mathbb{R}$.

**Proposition 7.** *Let $\vec{X} : S \to \mathbb{R}^n$ be a random variable.*

**(a)** *If $B \subset \mathbb{R}$ is an box, then $X^{-1}(B) \in \mathcal{E}$.*
**(b)** *If $\vec{x} \in \mathbb{R}^n$, then $\vec{X}^{-1}(x) \in \mathcal{E}$.*

**Remark 1.** Let $\{A_1, \ldots, A_n\}$ be a collection of sets and let $A = \times_{i=1}^n$ be their cartesian product. Define a function $\pi_i : A \to A_i$ by $\pi_i(a_1, \ldots, a_n) = a_i$. This function is called *projection on the* $i^{\text{th}}$ *component.*

Let $f : B \to A$ be a function. Define a function $f_i : B \to A_i$ by $f_i = \pi_i \circ f$. This function is called the $i^{\text{th}}$ *component function* of $f$. We see that $f(b) = (f_1(b), \ldots, f_n(b))$.

Let $\vec{a} = (a_1, \ldots, a_n) \in A$. Then $f^{-1}(\vec{a}) = \cap_{i=1}^n f_i^{-1}(a_i)$.

Let $A = A_1 \times A_2$. Let $f : B \to A$. Let $\vec{a} = (a_1, a_2)$. Then

**(a)** $f^{-1}(\vec{a}) = f_1^{-1}(a_1) \cap f_2^{-1}(a_2)$;
**(b)** $f_1^{-1}(a_1) = \cup_{a_2 \in \text{img}(f_2)} f_2^{-1}(a_2)$.

**Proposition 8.** *Let $\vec{X} : S \to \mathbb{R}^n$ and let $X_i : S \to \mathbb{R}$ be the $i^{\text{th}}$ component function of $\vec{X}$. Then $X_i$ is a random variable.*

**Definition 7.** Let $\vec{X} : S \to \mathbb{R}^n$ be a random vector.

We say that $\vec{X}$ is *discrete* if $\vec{X}(S)$ is countable.

**Definition 8.** Let $\vec{X} : S \to \mathbb{R}^n$ be a discrete random vector. The *joint density* of $\vec{X}$ is a function

$$f_{\vec{X}} : \mathbb{R} \to [0, 1] \text{ given by } f_{\vec{X}}(\vec{x}) = P(X^{-1}(\vec{x})).$$

**Proposition 9. Dirty Trick Theorem Revisited**
*Let $\vec{X} : S \to \mathbb{R}^n$ be a discrete random vector. Then*

$$\sum_{\vec{x} \in \text{img}(\vec{X})} f_{\vec{X}}(\vec{x}) = 1.$$

Let $[X = x]$ denote the preimage of $x$ under the random variable $X$.

**Proposition 10.** *Let $\vec{X} : S \to \mathbb{R}^n$ be a discrete random vector. Let $x \in \text{img}(\vec{X})$. Then $f_{\vec{X}}(x) = P(\cap_{i=1}^n [X_i = x_i])$.*

**Proposition 11.** *Let $\vec{X} : S \to \mathbb{R}^2$ be a discrete random vector. Let $X, Y : S \to \mathbb{R}$ be the components of $\vec{X}$. Then*

$$f_{X_1}(x) = \sum_{y \in \text{img}(Y)} f_{\vec{X}}(x, y).$$

**Multinomial Distribution**

Let $S$ be a finite set of cardinality $N$, and form the uniform probability space $(S, \mathcal{P}(S), P)$, where $P : \mathcal{P}(S) \to [0,1]$ is given by $P(E) = \frac{|E|}{|S|} = \frac{|E|}{N}$.

Let $R_1, \dots, R_n$ be disjoint events.

Let $R_0 = S \setminus \cup_{i=1}^n R_i$, so that $\{R_0, R_1, \dots, R_n\}$ form a partition of $S$.

Let $Y_0, Y_1, \dots, Y_n : S \to \mathbb{R}$ be the corresponding Bernoulli random variables.

Let $p_i = P(R_i)$.

Let $n$ be a positive integer. Let $T = \times_{i=1}^n S$, the cartesian product of $S$ with itself $n$ times. Then $|T| = N^n$. Form the uniform probability space $(T, \mathcal{P}(T), Q)$, where for $F \subset T$ we have $Q(F) = \frac{|Q|}{|T|} = \frac{|F|}{N^n}$.

Define discrete random vectors $X_i : T \to \mathbb{R}$ by

$$X(s_1, \dots, s_n) = \sum_{i=1}^n Y(s_i).$$

Define a discrete random vector $\vec{X} : T \to \mathbb{R}^n$ by $\vec{X} = (X_1, \dots, X_n)$.

**Multivariate Hypergeometric Distribution**

Let $S$ be a finite set of cardinality $N$, and form the uniform probability space $(S, \mathcal{P}(S), P)$, where $P : \mathcal{P}(S) \to [0,1]$ is given by $P(E) = \frac{|E|}{N}$.

Let $R_1, \dots, R_n$ be disjoint events.

Let $R_0 = S \setminus \cup_{i=1}^n R_i$, so that $\{R_0, R_1, \dots, R_n\}$ form a partition of $S$.

Let $Y_0, Y_1, \dots, Y_n : S \to \mathbb{R}$ be the corresponding Bernoulli random variables.

Let $p_i = P(R_i)$.

Let $n$ be an integer such that $0 \le n \le N$. Set

$$T = \{A \in \mathcal{P}(S) \mid |A| = n\}.$$

Then $|T| = \binom{N}{n}$. Form the uniform probability space $(T, \mathcal{P}(T), Q)$, where for $F \subset T$ we have $Q(F) = \frac{|F|}{|T|} = \frac{|F|}{\binom{N}{n}}$.

Define random variables $X_i : T \to \mathbb{R}$ by

$$X_i(A) = \sum_{a \in A} Y_i(a).$$

Then $X_i(A) = |A \cap R|$.

The image of $X$ is

$$\operatorname{img}(X) = \{0, 1, \dots, n\}.$$

The density of $X$ is

$$f_X(x) = \begin{cases} \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} & \text{if} \quad x \in \operatorname{img}(X); \\ 0 & \text{otherwise.} \end{cases}$$

The expectation of $X$ is

$$E(X) = \frac{nr}{N} = np.$$

Obtain this as follows. For $a \in S$, the number of sets in $T$ containing $a$ is $\binom{N-1}{n-1}$. Thus

$$\begin{aligned} E(X) &= \frac{1}{|T|} \sum_{A \in T} X(A) \\ &= \frac{1}{|T|} \sum_{A \in T} \sum_{a \in A} Y(a) \\ &= \frac{1}{|T|} \sum_{a \in R} |\{A \in T \mid a \in A\}| \\ &= \frac{1}{|T|} \sum_{a \in R} \binom{N-1}{n-1} \\ &= \frac{\binom{N-1}{n-1} r}{\binom{N}{n}} \\ &= \frac{nr}{N}. \end{aligned}$$

**Example 1.** An urn contains 2 red balls, three white balls, and four blue balls. One selects four balls at random from the urn without replacement. Let $X_1$ denote the number of red balls in the sample, let $X_2$ denote the number of white balls in the sample, and let $X_3$ denote the number of blue balls in the sample. Let $\vec{X} = (X_1, X_2, X_3)$.

- **(a)** Find the range of $(X, Y, Z)$.
- **(b)** Find the value of the joint density of $(X, Y, Z)$ at each point in the range.
- **(c)** Find the joint marginal density of $(X, Y)$, $(X, Z)$, and $(Y, Z)$.
- **(d)** Find the three univariate marginal densities.
- **(e)** Find the density of $X + Z$.
- **(f)** Find the expectations of $X$, $Y$, $Z$, $2X + 3Y$.

*Solution.* Let $S$ be the set of balls in the urn, together with the uniform probability structure.

The range is

$$\{(0,0,3),(0,1,2),(0,2,1),(0,3,0),(1,0,2),(1,1,1),(1,2,0),(2,0,1),(2,1,0)\}.$$

□

Department of Mathematics, Paragon Science Academy
*Email address*: pbailey@sonoranschools.org