Lecture 7

1 Common Distributions

1.1 Uniform Distribution – $\mathrm{U}(S)$

The Uniform distribution, parametrized with a *finite* set S, is a distribution μ over S, which is defined as $\mu(s) = \frac{1}{|S|}$ for every $s \in S$. Clearly this is a distribution as the support of μ is the finite set S, for every $s \in S$ we have $\mu(s) = \frac{1}{|S|} \in [0,1]$ and $\sum_{s\in S} \mu(s) = \sum_{s\in S} \frac{1}{|S|} = \frac{|S|}{|S|} = 1$.

When $S = \{a, \ldots, b\}$ is a set of consecutive integers, we will sometimes write $X \sim \mathrm{U}(a,b)$ instead of $X \sim \mathrm{U}(\{a, a+1, \ldots, b\})$.

Example 1: Let X be the outcome of one roll of a fair die. As we previously saw $X \sim \mathrm{U}(1,6)$.

1.2 Bernoulli Distribution – $\mathrm{Ber}(p)$

The Bernoulli distribution, parametrized with a real number $0 \le p \le 1$, is a distribution μ over $\{0,1\}$, which is defined as $\mu(1) = p$ and $\mu(0) = 1-p$. Clearly this is a distribution as the support of μ is the finite set $\{0,1\}$, for every $s \in \{0,1\}$ we have $0 \le \mu(s) \le 1$ and $p + (1-p) = 1$.

Example 2: Let A be an event with probability p. Then, as we previously saw, 1_A is random variable which satisfies $1_A \sim \mathrm{Ber}(p)$. For example, toss a coin with probability p for heads. Let A be the event "the outcome of the coin toss is heads". Then $1_A \sim \mathrm{Ber}(p)$.

1.3 Binomial Distribution – $\mathrm{Bin}(n,p)$

The Binomial distribution, parametrized with a natural number n and a real number $0 \le p \le 1$, is a distribution μ over $\{0,1,\ldots,n\}$, which is defined as

$$\mu(k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

for every $k \in \{0,1,\ldots,n\}$. This is a distribution as the support of μ is the finite set $\{0,1,\ldots,n\}$, and since, moreover, $\mu(k) \ge 0$ for every k, it suffices to show that $\sum_k \mu(k) = 1$. Indeed

$$\sum_k \mu(k) = \sum_{k=0}^{n} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} = (p + (1-p))^n = 1,$$

where the second equality holds by the binomial formula.

Example 3: Toss a coin with probability p for heads n times, all coin tosses being mutually independent. Let X be the total number of tosses whose outcome was heads. Then $X \sim \mathrm{Bin}(n, p)$.

The following theorem relates the Binomial and Bernoulli distributions (and also justifies the example above). Before stating it, we need the following definition.

Definition 1.1 (Independence of Random Variables)**.** *For every $1 \leq i \leq n$, let $X_i : \Omega \to S_i$ be a random variable. $X_1, \ldots, X_n$ are said to be* independent *(or* mutually independent*) if*

$$\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n) = \mathbb{P}(X_1 = x_1) \cdot \ldots \cdot \mathbb{P}(X_n = x_n)$$

holds for every $(x_1, \ldots, x_n) \in S_1 \times \ldots \times S_n$.

Theorem 1.2. *Let $n \in \mathbb{N}$ and let $0 \leq p \leq 1$ be a real number. Let $X_1, X_2, \ldots, X_n \sim \mathrm{Ber}(p)$ be independent random variables, and let $S_n = X_1 + X_2 + \ldots + X_n$. Then $S_n \sim \mathrm{Bin}(n, p)$.*

Proof. Since every X_i takes values in $\{0, 1\}$, it follows that S_n takes values in $\{0, 1, \ldots, n\}$. For any $\mathcal{I} \subseteq \{1, \ldots, n\}$ define the event

$$A_{\mathcal{I}} = \left(\bigcap_{i \in \mathcal{I}} \{X_i = 1\} \right) \cap \left(\bigcap_{i \notin \mathcal{I}} \{X_i = 0\} \right).$$

Then for every $k \in \{0, 1, \ldots, n\}$ it holds that

$$\{S_n = k\} = \bigcup_{\mathcal{I} : |\mathcal{I}| = k} A_{\mathcal{I}}.$$

Since these events are pairwise disjoint we get that

$$\begin{aligned}
\mathbb{P}(S_n = k) &= \mathbb{P}\left(\bigcup_{\mathcal{I} : |\mathcal{I}| = k} A_{\mathcal{I}} \right) \\
&= \sum_{\mathcal{I} : |\mathcal{I}| = k} \mathbb{P}(A_{\mathcal{I}}) \\
&= \sum_{\mathcal{I} : |\mathcal{I}| = k} \prod_{i \in \mathcal{I}} \mathbb{P}(X_i = 1) \cdot \prod_{i \notin \mathcal{I}} \mathbb{P}(X_i = 0) \\
&= \sum_{\mathcal{I} : |\mathcal{I}| = k} p^k \cdot (1 - p)^{n-k} \\
&= \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k},
\end{aligned}$$

where the third equality holds since $X_1, X_2, \ldots, X_n$ are independent and the fourth equality holds by the definition of the Bernoulli distribution. $\square$

1.4 Geometric Distribution – $\mathrm{Geom}\,(p)$

The Geometric distribution, parametrized with a real number $0 < p < 1$, is a distribution μ over $\mathbb{N}$ (without 0), which is defined as

$$\mu\,(k) = (1-p)^{k-1} \cdot p,$$

for every $k \in \mathbb{N}$. This is a distribution as the support of μ is the countably infinite set $\mathbb{N}$, and since, moreover, $\mu(k) \geq 0$ for every k, it suffices to show that $\sum_k \mu(k) = 1$. Indeed, by the formula for the sum of an infinite geometric series, we have

$$\sum_k \mu(k) = \sum_{k=1}^{\infty} (1-p)^{k-1} \cdot p = p \cdot \frac{1}{1-(1-p)} = 1.$$

Example 4: Toss a coin with probability p for heads until the first time the outcome is heads, all coin tosses being mutually independent. Let X be the total number of tosses. Then $X \sim \mathrm{Geom}\,(p)$. Indeed, if $X = k$, then in each of the first $k-1$ tosses the outcome was tails and in the kth toss it was heads. Since all coin tosses are mutually independent, this happens with probability $(1-p)^{k-1} \cdot p$.

1.5 Hypergeometric Distribution – $\mathrm{Hyp}\,(N, D, n)$

The Hypergeometric distribution, parametrized with 3 natural numbers N, D, and n, is a distribution μ over $\{\max\{0, n+D-N\}, \ldots, \min\{n, D\}\}$, which is defined as

$$\mu\,(k) = \frac{\binom{D}{k} \cdot \binom{N-D}{n-k}}{\binom{N}{n}}$$

for every $k \in \{0, 1, \ldots, n\}$. The proof that μ is a distribution is left as an exercise.

Example 5: Consider an urn that contains N balls of which D are red and $N-D$ are blue. We draw n balls uniformly at random from the urn *without* replacement. Let X be the number of red balls that were drawn from the urn. Then $X \sim \mathrm{Hyp}\,(N, D, n)$. Indeed, since the balls are drawn uniformly without replacement, it follows that

$$\mathbb{P}\,(X = k) = \frac{|\{X = k\}|}{\binom{N}{n}}$$

for every $0 \leq k \leq n$. Moreover, every choice of k red balls and $n-k$ blue balls is counted exactly once in $\{X = k\}$. Hence $|\,\{X = k\}\,| = \binom{D}{k} \cdot \binom{N-D}{n-k}$. We conclude that

$$\mathbb{P}\,(X = k) = \frac{\binom{D}{k} \cdot \binom{N-D}{n-k}}{\binom{N}{n}}$$

as claimed

Remark 1.3. *If the balls are drawn* ***with*** *replacement (and still independently and uniformly at random), then* $X \sim \mathrm{Bin}\,(n, D/N)$. *Whenever* n *is "much smaller" than* N *and* D*, the binomial distribution* $\mathrm{Bin}\,(n, D/N)$ *is a good approximation of the hypergeometric distribution* $\mathrm{Hyp}\,(N, D, n)$.