# Random Variables

Stephen Koo

June 23, 2015

---

# 1 Discrete Random Variables

## 1.1 Bernoulli

An experiment that results in "success" or "failure."

$$\begin{aligned} X &\sim \mathrm{Ber}(p) \\ P(X=0) &= 1-p \\ P(X=1) &= p \\ E[X] &= p \\ \mathrm{Var}(X) &= p(1-p) \\ M(t) &= e^t p + 1 - p \end{aligned}$$

## 1.2 Binomial

The number of successes in an experiment with $n$ trials and $p$ probability of success on each trial.

$$\begin{aligned} X &\sim \mathrm{Bin}(n,p) \\ P(X=i) = p(i) &= \binom{n}{i} p^i (1-p)^{n-i} \quad \text{where } i = 0, 1, \ldots, n \\ E[X] &= np \\ \mathrm{Var}(X) &= np(1-p) \\ M(t) &= (pe^t + 1 - p)^n \end{aligned}$$

If $X_i \sim \mathrm{Bin}(n_i, p)$ for $1 \le i \le N$, then

$$\left(\sum_{i=1}^{N} X_i\right) \sim \mathrm{Bin}\left(\sum_{i=1}^{N} n_i, p\right)$$

Note that the binomial distribution is a generalization of the Bernoulli distribution, since $\mathrm{Ber}(p) \sim \mathrm{Bin}(1, p)$.

## 1.3 Poisson

Approximates the binomial random variable when $n$ is large and $p$ is small enough to make $np$ "moderate"—generally when $n > 20$ and $p < 0.05$—and approaches the binomial distribution as $n \to \infty$ and $p \to 0$.

$$\begin{aligned} X &\sim \text{Poi}(\lambda) \quad \text{where } \lambda = np \\ P(X = i) &= e^{-\lambda}\frac{\lambda^i}{i!} \quad \text{where } i = 0, 1, 2, \ldots \\ E[X] &= \lambda \\ \text{Var}(X) &= \lambda \\ M(t) &= e^{\lambda(e^t - 1)} \end{aligned}$$

The approximations also works to a certain extent when the successes in the trials are not entirely independent, and when the probability of success in each trial varies slightly.

If $X_i \sim \text{Poi}(\lambda_i)$ for $1 \le i \le N$, then

$$\left(\sum_{i=1}^{N} X_i\right) \sim \text{Poi}\left(\sum_{i=1}^{N} \lambda_i\right)$$

## 1.4 Geometric

The number of independent trials until a success, where the probability of success is $p$.

$$\begin{aligned} X &\sim \text{Geo}(p) \\ P(X = n) &= (1-p)^{n-1}p \quad \text{where } n = 1, 2, \ldots \\ E[X] &= 1/p \\ \text{Var}(X) &= (1-p)/p^2 \end{aligned}$$

## 1.5 Negative Binomial

The number of independent trials until $r$ successes, with probability $p$ of success.

$$\begin{aligned} X &\sim \text{NegBin}(r, p) \\ P(X = n) &= \binom{n-1}{r-1} p^r (1-p)^{n-r} \quad \text{where } n = r, r+1, \ldots \\ E[X] &= r/p \\ \text{Var}(x) &= r(1-p)/p^2 \\ \text{Geo}(p) &\sim \text{NegBin}(1, p) \end{aligned}$$

Note that the negative binomial distribution generalizes the geometric distribution, with $\text{Geo}(p) \sim \text{NegBin}(1, p)$.

### 1.6 Hypergeometric

The number of white balls drawn after drawing $n$ balls (without replacement) from an urn containing $N$ balls, with $m$ white balls and $N - m$ other ("black") balls.

$$X \sim \text{HypG}(n, N, m)$$
$$P(X = i) = \frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}} \quad \text{where } i = 0, 1, \ldots, n$$
$$E[X] = n(m/N)$$
$$\text{Var}(X) = \frac{nm(N-n)(N-m)}{N^2(N-1)}$$
$$\text{HypG}(n, N, m) \to Bin(n, m/N) \text{ , as } N \to \infty \text{ and } m/N \text{ stays constant}$$

### 1.7 Multinomial

The multinomial distribution further generalizes the binomial distribution: given an experiment with $n$ independent trials, where each trial results in one of $m$ outcomes, with respective probabilities $p_1, p_2, \ldots, p_m$ such that $\sum_{i=1}^{m} p_i = 1$, then if $X_i$ denotes the number of trials with outcome $i$ we have

$$P(X_1 = c_1, X_2 = c_2, \ldots, X_m = c_m) = \binom{n}{c_1, c_2, \ldots, c_m} p_1^{c_1} p_2^{c_2} \cdots p_m^{c_m}$$

where $\sum_{i=1}^{m} c_i = n$ and $\binom{n}{c_1, c_2, \ldots, c_m} = \frac{n!}{c_1! c_2! \cdots c_m!}$.

## 2 Continuous Random Variables

If $Y$ is a non-negative continuous random variable

$$E[Y] = \int_0^\infty P(Y > y) dy$$

### 2.1 Uniform

$$X \sim \text{Uni}(\alpha, \beta)$$
$$f(x) = \begin{cases} \frac{1}{\beta - \alpha} & \alpha \le x \le \beta \\ 0 & \text{otherwise} \end{cases}$$
$$E[X] = \frac{\alpha + \beta}{2}$$
$$\text{Var}(X) = \frac{(\beta - \alpha)^2}{12}$$

## 2.2 Normal

For values in common natural phenomena, especially when resulting from the sum of multiple variables.

$$
\begin{aligned}
X &\sim \mathrm{N}(\mu, \sigma^2) \\
f(x) &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{where } -\infty < x < \infty \\
E[X] &= \mu \\
\mathrm{Var}(X) &= \sigma^2 \\
M(t) &= e^{\left(\frac{\sigma^2 t^2}{2} + \mu t\right)}
\end{aligned}
$$

Letting $X \sim N(\mu, \sigma^2)$ and $Y = aX + b$, we have

$$
\begin{aligned}
Y &\sim \mathrm{N}(a\mu + b, a^2\sigma^2) \\
F_Y(x) &= F_X(\frac{x-b}{a})
\end{aligned}
$$

The Standard (Unit) Normal Random Variable $Z \sim N(0,1)$ has a cumulative distribution function (CDF) commonly labeled $\Phi(z) = P(Z \leq z)$ that has some useful properties.

$$
\begin{aligned}
\Phi(z) &= \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \\
\Phi(-z) &= 1 - \Phi(z) \\
P(Z \geq -z) &= P(Z > z)
\end{aligned}
$$

Given $X \sim N(\mu, \sigma^2)$ where $\sigma > 0$, we can then compute the CDF of $X$ using the CDF of the standard normal variable.

$$F_X(x) = \Phi(\frac{x-\mu}{\sigma})$$

By the de Moivre-Laplace Limit Theorem, the normal variable can approximate the binomial when $\mathrm{Var}(X) = np(1-p) \geq 10$. If we let $S_n$ denote the number of successes (with probability $p$) in $n$ independent trials, then

$$P\left(a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\right) \overset{n\to\infty}{\to} \Phi(b) - \Phi(a)$$

If $X_i \sim \mathrm{N}(\mu_i, \sigma_i^2)$ for $i = 1, 2, \ldots, n$, then

$$\left(\sum_{i=1}^{n} X_i\right) \sim \mathrm{N}\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right)$$

## 2.3 Exponential

Represents time until some event, with rate $\lambda > 0$.

$$
\begin{aligned}
X &\sim \mathrm{Exp}(\lambda) \\
f(x) &= \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \\
E[X] &= \frac{1}{\lambda} \\
\mathrm{Var}(X) &= \frac{1}{\lambda^2} \\
F(x) &= 1 - e^{-\lambda x} \quad \text{where } x \geq 0
\end{aligned}
$$

Exponentially distributed random variables are memoryless.

$$P(X > s + t | X > s) = P(X > t)$$

## 2.4 Beta

$$
\begin{aligned}
X &\sim \mathrm{Beta}(a, b) \\
f(x) &= \begin{cases} \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1} & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases} \\
B(a, b) &= \int_0^1 x^{a-1}(1-x)^{b-1} dx \\
E[X] &= \frac{a}{a+b} \\
\mathrm{Var}(X) &= \frac{ab}{(a+b)^2(a+b+1)}
\end{aligned}
$$

If $X \sim \mathrm{Uni}(0, 1)$ and $N$ denotes the number of heads resulting from a number of coin flips with some unknown probability of getting heads, then

$$X | (N = n, m + n \text{ trials}) \sim \mathrm{Beta}(n + 1, m + 1)$$