Some special distributions

- For a specific univariate distribution, we often characterize it using its PMF/PDF/MGF and find the mean and variance of the distribution. The Chebyshev's inequality gives us an approximate range of a random variable X based on $\mu = E(X)$ and $\sigma^2 = Var(X)$:

$$P(X \in (\mu - k\sigma, \mu + k\sigma)) \geq 1 - \frac{1}{k^2}$$

for $k > 0$.

- Bernoulli trial. A Bernoulli trial is an experiment of two possible outcomes: success and failure, denoted by 1 and 0 respectively. $P(\text{outcome} = 1)$ is called the success probability.

- Binomial distribution (二項分布) $b(n, p)$. Suppose that we have results of n independent Bernulli trials with success probability p. Let X be the number of successes in the n trials. Then the distribution of X is the binomial distribution with size n and success probability p, denoted by $b(n, p)$. The PMF of X is p_X, where

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

for $x \in \{0, 1, \ldots, n\}$.

  - The binomial distribution $b(1, p)$ is also called the Bernoulli distribution with success probability p.

- Notation. For a distribution $\mathcal{D}$, $P(\mathcal{D} \in A)$ means $P(X \in A)$, where $X \sim \mathcal{D}$. For example,

  - $P(N(0, 1) \leq x)$ means $P(Z \leq x)$, where $Z \sim N(0, 1)$.
  - $P(b(10, 0.5) = x)$ means $P(X = x)$, where $X \sim b(10, 0.5)$.

- Example 1. Find the mean and variance of $b(n, p)$ by finding its MGF.

  A sketch of solution. Let $M_{n,p}$ be the MGF of $b(n, p)$, then $M_{n,p}$ is the MGF for a random variable X such that $X \sim b(n, p)$, so for $t \in (-\infty, \infty)$,

$$\begin{aligned} M_{n,p}(t) &= \sum_{x=0}^{n} e^{tx} P(b(n, p) = x) \\ &= (1 - p + pe^t)^n. \end{aligned}$$

  Compute $M'_{n,p}(t)$ and $M''_{n,p}(t)$, then we have $M'_{n,p}(0) = np$ and $M''_{n,p}(0) = n(n-1)p^2 + np$, so the mean of $b(n, p)$ is np and the variance of $b(n, p)$ is $n(n-1)p^2 + np - (np)^2 = np(1 - p)$.

- Fact 1 Suppose that X_1, …, X_n are IID and $X_1 \sim b(1, p)$, then $\sum_{i=1}^n X_i \sim b(n, p)$.
- Negative binomial distribution (負二項分布) $nb(r, p)$. Suppose that we continue to run independent Bernulli trials with success probability p until r successes are obtained. Let X be the number of failures before the r-th success. Then the disitribution of X is a negative binomial distribution, denoted by $nb(r, p)$.

  $$P(nb(r, p) = x) = C_x^{x+r-1}(1-p)^x p^r$$

  for $x \in \{0, 1, \ldots\}$.

  - When $r = 1$, $nb(1, p)$ is called a geometric distribution.
- Multinomial distribution. Suppose that X is a random variable that takes values in $\{1, \ldots, k\}$ and $P(X = j) = p_j$ for $j = 1$, …, k. Suppose that X_1, …, X_n are IID and $X_1 \sim X$. Let

  $$N_j = \sum_{i=1}^{n} I(X_i = j)$$

  for $j = 1$, …, k, where

  $$I(X_i = j) = \begin{cases} 1 & \text{if } X_i = j; \\ 0 & \text{if } X_i \neq j, \end{cases}$$

  then the distribution of $(N_1, \ldots, N_k)$ is called a multinomial distribution of size n and probability vector $(p_1, \ldots, p_k)$.

  $$P((N_1, \ldots, N_k) = (x_1, \ldots, x_k)) = \frac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k}$$

  for x_1, …, $x_k \in \{0, \ldots, n\}$ and $\sum_{j=1}^n x_k = n$.
- Hypergeometric distribution (超幾何分布) $H(N, S, n)$. Suppose that we have a group of N items, of which S items are good and $N - S$ items are defective. Choose a sample of n items from the group and let X be the number of good items in the sample. Then the distribution of X, is a hypergeometric distribution, denoted by $H(N, S, n)$.

  $$P(H(N, S, n) = x) = \frac{C_x^S C_{n-x}^{N-S}}{C_n^N}$$

  if $x \in \{0, 1, \ldots, n\}$, $x \leq S$ and $n - x \leq N - S$.

- Poisson distribution $Poisson(\mu)$. For a constant $\mu > 0$, the Poisson distribution with mean μ, denoted by $Poisson(\mu)$, is the distribution with PMF p_μ, where

$$p_\mu(x) = \frac{e^{-\mu}\mu^x}{x!} \text{ for } x \in \{0, 1, \ldots\}.$$

  - From the solution to Problem 30, the mean of $Poisson(\lambda)$ is λ.
  - The MGF of $Poisson(\lambda)$ and the second moment of $Poisson(\lambda)$ can be also found in the solution to Problem 30.
  - $Poisson(\mu)$ is the limit of $b(n, \mu/n)$ as $n \to \infty$ in the sense that

$$\lim_{n\to\infty} P(b(n, \mu/n) = x) = \frac{e^{-\mu}\mu^x}{x!} = P(Poisson(\mu) = x)$$

    for $x \in \{0, 1, \ldots\}$.
  - The MGF of $Poisson(\lambda)$ can be found in the solution to Problem 30. From the solution to Problem 30, we also have the first moment and second moment of $Poisson(\lambda)$ are λ and $\lambda + \lambda^2$ respectively, so the mean and variance of $Poisson(\lambda)$ are both equal to λ.

- Poisson processes. Suppose that for $t \geq 0$, $N(t)$ denotes the number of times that certain event occured in the time interval $[0, t]$ and $N(0) = 0$. $\{N(t)\}_{t\geq 0}$ is called a Poisson process with rate parameter λ ($\lambda > 0$) if the statements in (a) and (b) hold.

  (a) For $t \geq 0$ and $h > 0$, $N(t + h) - N(t) \sim Possion(\lambda h)$.

  (b) For t_1, $t_2 \geq 0$ and h_1, $h_2 > 0$, if $(t_1, t_1+h_1) \cap (t_2, t_2+h_2)$ is an empty set, then $N(t_1+h_1) - N(t_1)$ and $N(t_2+h_2) - N(t_2)$ are independent.

- Fact 2 Suppose that $\{N(t)\}_{t\geq 0}$ is a Poisson process with rate parameter λ. Let W_k be the time for the k-th occurrence of the process event for $k = 1, 2, \ldots$. Then the interarrival times W_1, $W_2 - W_1$, $W_3 - W_2$, $\ldots$ are IID and

$$P(W_1 \leq t) = \begin{cases} 1 - P(N(t) = 0) = 1 - e^{-\lambda t} & \text{if } t > 0; \\ 0 & \text{if } t \leq 0. \end{cases} \tag{1}$$

- The distribution of W_1 in (1) is called the exponential distribution with mean $1/\lambda$. A PDF of the exponential distribution with mean $1/\lambda$ is f, where

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x > 0; \\ 0 & \text{if } x \leq 0. \end{cases}$$

- Gamma distribution $\Gamma(\alpha, 1)$. Suppose that $\alpha > 0$. Define

$$f(x) = \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} I_{(0,\infty)}(x)$$

for $x \in R$, where

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx.$$

Then the distribution with PDF f is called the gamma distribution with shape parameter α and scale parameter 1, denoted by $\Gamma(\alpha, 1)$ in the textbook.

  - $\Gamma(1, 1)$ is the exponential distribution with mean 1.
  - Suppose that $X \sim \Gamma(\alpha, 1)$, then the distribution of βX is also a gamma distribution.

- Fact 3 Suppose that $T_1, \ldots, T_k$ are IID random variables and the distribution of T_1 is the exponential distirbution with mean $1/\lambda$. Then the distribuion of $\lambda \sum_{i=1}^k T_i$ is the gamma distribution $\Gamma(k, 1)$.

- Definition. For $r > 0$, The chi-square distribution with degrees of freedom r, denoted by $\chi^2(r)$, is the distribution of $2X$, where $X \sim \Gamma(r/2, 1)$. When r is a positive integer, let $Z_1, \ldots, Z_r$ be IID random variables such that $Z_1 \sim N(0, 1)$, then $\sum_{i=1}^r Z_i^2 \sim \chi^2(r)$.

- Definition. Suppose that $a > 0$, $b > 0$, $X_1 \sim \Gamma(a, 1)$, $X_2 \sim \Gamma(b, 1)$ and X_1 and X_2 are indepdendent. Then the distrbution of $X_1/(X_1 + X_2)$ is the beta (β) distribution with PDF f, where

$$f(x) = \begin{cases} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} & \text{if } x \in (0, 1); \\ 0 & \text{if } x \notin (0, 1), \end{cases}$$

- Definition. Suppose that $r > 0$, $Z \sim N(0, 1)$, $V \sim \chi^2(r)$ and Z and V are independent. Then the distribution of

$$\frac{Z}{\sqrt{(V/r)}}$$

is the t distribution with r degrees of freedom, denoted by $t(r)$.

- Definition. Suppose that $d > 0$, $r > 0$, $U \sim \chi^2(d)$, $V \sim \chi^2(r)$ and U and V are independent. Then the distribution of

$$\frac{(U/d)}{(V/r)}$$

is the F distribution with degrees of freedom d and r, denoted by $F(d, r)$.