Common Families of Distributions

Discrete Distributions

Discrete Uniform Distribution

A random variable X has a discrete uniform $(1, N)$ distribution if

$$P(X = x \mid N) = \frac{1}{N}, \quad x = 1, 2, \ldots, N$$

$$\mathrm{E}X = \frac{N+1}{2}$$

$$\mathrm{E}X^2 = \sum_{x=1}^{N} x^2 \frac{1}{N} = \frac{(N+1)(2N+1)}{6}$$

$$\operatorname{Var} X = \frac{(N+1)(N-1)}{12}$$

Hypergeometric Distribution

N balls, M red, N-M green. take K balls, x red.

$$P(X = x \mid N, M, K) = \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}}, \quad x = 0, 1, \ldots, K$$

$$M - (N - K) \leq x \leq M$$

$$\mathrm{EX} = \frac{KM}{N}$$

$$\mathrm{Var}\, X = \frac{KM}{N}\left(\frac{(N-M)(N-K)}{N(N-1)}\right)$$

Binomial Distribution (Count Success)

Counts the number of successes in a fixed number of Bernoulli trials.

$$P(Y=y \mid n,p) = \binom{n}{y} p^y (1-p)^{n-y}, \quad y = 0,1,2,\ldots,n,$$

Theorem 3.2.2 (Binomial Theorem) For any real numbers x and y and integer $n \geq 0$,

$$(x+y)^n = \sum_{=0}^{n} \binom{n}{i} x^n y^{n-i}$$

$$2^n = \sum_{i=0}^{n} \binom{n}{i}$$

$$\mathrm{EX} = np, \quad \mathrm{Var}\, X = np(1-p)$$

$$M_X(t) = \left[pe^t + (1-p)\right]^n$$

Poisson Distribution (Waiting Time)

For small time intervals, the probability of an arrival is proportional to the length of waiting time.

λ is intensity parameter.

$$P(X=x \mid \lambda) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0,1,\ldots$$

$$e^y = \sum_{i=0}^{\infty} \frac{y^i}{i!}$$

Thus,

$$\sum_{x=0}^{\infty} P(X=x \mid \lambda) = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{-\lambda} e^{\lambda} = 1$$

The mean of X is easily seen to be

$$\begin{aligned}
\mathrm{EX} &= \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} \\
&= \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} \\
&= \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} \\
&= \lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} \qquad \text{(substitute } y = x-1\text{)}
\end{aligned}$$

A similar calculation will show that

$$\operatorname{Var} X = \lambda$$

and so the parameter λ is both the mean and the variance of the Poisson distribution.

$$M_X(t) = e^{\lambda(e^t - 1)}$$

Negative Binomial Distribution (Wait for r Successes)

Count the number of Bernoulli trials required to get a fixed number of successes.

$$P(X = x \mid r, p) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \quad x = r, r+1, \ldots$$

The negative binomial distribution is sometimes defined in terms of the random variable $Y =$ number of failures before the r th success.

$$P(Y = y) = \binom{r+y-1}{y} p^r (1-p)^y, \quad y = 0, 1, \ldots.$$

$$\binom{r+y-1}{y} = (-1)^y \binom{-r}{y} = (-1)^y \frac{(-r)(-r-1)(-r-2)\cdot\cdots\cdots(-r-y+1)}{(y)(y-1)(y-2)\cdot\cdots\cdot(2)(1)}$$

$$P(Y = y) = (-1)^y \binom{-r}{y} p^r (1-p)^y$$

a striking resemblance to the binomial distribution.

$$\mathrm{E}\mathrm{Y} = r\frac{(1-p)}{p}.$$

$$\operatorname{Var} Y = \frac{r(1-p)}{p^2}$$

If we define the parameter $\mu = r(1-p)/p$, then

$$\operatorname{Var} Y = \mu + \frac{1}{r}\mu^2$$

The variance is a quadratic function of the mean.

The negative binomial family of distributions includes the Poisson distribution as a limiting case. If $r \to \infty$ and $p \to 1$ such that $r(1-p) \to \lambda, 0 < \lambda < \infty$, then

$$\mathrm{E} Y = \frac{r(1-p)}{p} \to \lambda$$

$$\operatorname{Var} Y = \frac{r(1-p)}{p^2} \to \lambda$$

Geometric Distribution (Wait for one success)

A special case of the negative binomial distribution. If we set $r = 1$ we have

$$P(X = x \mid p) = p(1-p)^{x-1}, \quad x = 1, 2, \ldots$$

The mean and variance of X can be calculated by using the negative binomial formulas and by writing $X = Y + 1$ to obtain

$$\mathrm{E}X = \frac{1}{p} \quad \text{and} \quad \mathrm{Var}\, X = \frac{q}{p^2}$$

The geometric distribution "memoryless" property. For integers $s > t$, it is the case that

$$P(X > s \mid X > t) = P(X > s - t)$$

The probability of getting an additional $s - t$ failures, having already observed t failures, is the same as the probability of observing $s - t$ failures at the start of the sequence.

$$\begin{aligned} P(X > n) &= P(\text{ no successes in } n \text{ trials }) \\ &= (1-p)^n \end{aligned}$$

and hence

$$\begin{aligned} P(X > s \mid X > t) &= \frac{P(X > s \text{ and } X > t)}{P(X > t)} \\ &= \frac{P(X > s)}{P(X > t)} \\ &= (1-p)^{s-t} \\ &= P(X > s - t) \end{aligned}$$

Continuous Distributions

Uniform Distribution

$$f(x \mid a, b) = \begin{cases} \dfrac{1}{b-a} & \text{if } x \in [a, b] \\ \text{otherwise} \end{cases}$$

$$\begin{aligned} \mathrm{EX} &= \int_a^b \frac{x}{b-a} dx = \frac{b+a}{2} \\ \mathrm{Var}\, X &= \int_a^b \frac{\left(x - \frac{b+a}{2}\right)^2}{b-a} dx = \frac{(b-a)^2}{12} \end{aligned}$$

Gamma Distribution

If α is a positive constant,

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$$

The gamma function satisfies many useful relationships, in particular

$$\Gamma(\alpha+1)=\alpha\Gamma(\alpha), \quad \alpha>0$$

which can be verified through integration by parts. $\Gamma(1)=1$, we have for any integer $n>0$,

$$\Gamma(n)=(n-1)!$$

(Another useful special case is that $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$.)

$$f(t)=\frac{t^{\alpha-1}e^{-t}}{\Gamma(\alpha)}, \quad 0<t<\infty$$

is a pdf. The full gamma family, however, has two parameters, and can be derived by changing variables to get the pdf of the random variable $X=\beta T$.

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

α: shape parameter, peakedness of the distribution

β: scale parameter, spread of the

$$\begin{aligned}
\mathrm{EX} &= \frac{1}{\Gamma(\alpha)\beta^{\alpha}}\int_0^{\infty} x^{\alpha}e^{-x/\beta}dx \\
&= \frac{1}{\Gamma(\alpha)\beta^{\alpha}}\Gamma(\alpha+1)\beta^{\alpha+1} \\
&= \frac{\alpha\Gamma(\alpha)\beta}{\Gamma(\alpha)} \\
&= \alpha\beta
\end{aligned}$$

$$\operatorname{Var} X=\alpha\beta^2$$

$$M_X(t)=\left(\frac{1}{1-\beta t}\right)^{\alpha}, \quad t<\frac{1}{\beta}$$

Example 3.3.1 (Gamma-Poisson relationship) If X is a gamma (α,β) random variable, where α is an integer, then for any x,

$$P(X\le x)=P(Y\ge\alpha)$$

where $Y\sim$ Poisson (x/β). (established by successive integrations by parts, as follows. Since α is an integer, we write $\Gamma(\alpha)=(\alpha-1)!$)

Chi Squared (p/2, 2)

If we set $\alpha=p/2$. where p is an integer, and $\beta=2$, then the gamma pdf becomes

$$f(x \mid p)=\frac{1}{\Gamma(p/2)2^{p/2}}x^{(p/2)-1}e^{-x/2}, \quad 0<x<\infty$$

which is the chi squared pdf with p degrees of freedom. The mean, variance, and mgf of the chi squared distribution can all be calculated by using the previously derived gamma formulas.

Exponential Distribution (1)

When we set $\alpha = 1$. We then have

$$f(x \mid \beta) = \frac{1}{\beta} e^{-x/\beta}, \quad 0 < x < \infty,$$

the exponential pdf with scale parameter β. The exponential distribution can be used to model lifetimes, analogous to the use of the geometric distribution in the discrete case. In fact, the exponential distribution shares the "memoryless" property of the geometric. If $X \sim \text{exponential}(\beta)$, then for $s > t \geq 0$,

$$\begin{aligned} P(X > s \mid X > t) &= \frac{P(X > s, X > t)}{P(X > t)} \\ &= \frac{P(X > s)}{P(X > t)} \\ &= \frac{\int_s^\infty \frac{1}{\beta} e^{-x/\beta} dx}{\int_t^\infty \frac{1}{\beta} e^{-x/\beta} dx} \\ &= \frac{e^{-s/\beta}}{e^{-t/\beta}} \\ &= e^{-(s-t)/\beta} \\ &= P(X > s - t) \end{aligned}$$

Normal distribution

$$f\left(x \mid \mu, \sigma^2\right) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}, \quad -\infty < x < \infty$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} dz = 1$$

$$\int_0^\infty e^{-z^2/2} dz = \frac{\sqrt{2\pi}}{2} = \sqrt{\frac{\pi}{2}}$$

$$t = r\cos\theta \quad \text{ and } \quad u = r\sin\theta.$$

Then $r^2 + u^2 = r^2$ and $dtdu = rd\theta dr$ and the limits of integration become $0 < r < \infty, 0 < \theta < \pi/2$ (the upper limit on θ is $\pi/2$ because t and u are restricted to be positive). We now have

$$\begin{aligned} \int_0^\infty \int_0^\infty e^{-(r^2+u^2)/2} dtdu &= \int_0^\infty \int_0^{\pi/2} re^{-r^2/2} d\theta dr \\ &= \frac{\pi}{2} \int_0^\infty re^{-r^2/2} dr \\ &= \frac{\pi}{2} \left[-e^{-r^2/2} \Big|_0^\infty \right] \\ &= \frac{\pi}{2} \end{aligned}$$

Making the substitution $w = \frac{1}{2} z^2$,

$$\Gamma\left(\frac{1}{2}\right) = \int_0^\infty w^{-1/2} e^{-w} dw = \sqrt{\pi}$$

$$\begin{aligned} P(|X-\mu| \le \sigma) = P(|Z| \le 1) = .6826 \\ P(|X-\mu| \le 2\sigma) = P(|Z| \le 2) = .9544 \\ P(|X-\mu| \le 3\sigma) = P(|Z| \le 3) = .9974 \end{aligned}$$

Beta Distribution

$$f(x \mid \alpha, \beta) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1, \quad \alpha > 0, \quad \beta > 0$$

where $B(\alpha,\beta)$ denotes the beta function,

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx$$

The beta function is related to the gamma function through the following identity:

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

For $n > -\alpha$ we have

$$\begin{aligned} EX^n &= \frac{1}{B(\alpha,\beta)} \int_0^1 x^n x^{\alpha-1}(1-x)^{\beta-1} dx \\ &= \frac{1}{B(\alpha,\beta)} \int_0^1 x^{(\alpha+n)-1}(1-x)^{\beta-1} dx \end{aligned}$$

We now recognize the integrand as the kernel of a beta $(\alpha+n,\beta)$ pdf, hence

$$\mathrm{E} X^n = \frac{B(\alpha+n,\beta)}{B(\alpha,\beta)} = \frac{\Gamma(\alpha+n)\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+n)\Gamma(\alpha)}$$

$$\mathrm{E}\mathrm{X} = \frac{\alpha}{\alpha+\beta} \quad \text{and} \quad \operatorname{Var} X = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

Cauchy Distribution

$$f(x \mid \theta) = \frac{1}{\pi} \frac{1}{1+(x-\theta)^2}, \quad -\infty < x < \infty, \quad -\infty < \theta < \infty,$$

A surprising fact is that the ratio of two standard normals has a Cauchy distribution. Taking ratios can lead to ill-behaved distributions.

Lognormal Distribution

$\log X \sim \mathrm{n}\left(\mu, \sigma^2\right)$),

$$\begin{aligned} \mathrm{EX} &= \mathrm{E} e^{\log X} \\ &= \mathrm{E} e^{Y} \qquad \left(Y = \log X \sim \mathrm{n}\left(\mu, \sigma^2\right)\right) \\ &= e^{\mu + \left(\sigma^2/2\right)} \end{aligned}$$

We can use a similar technique to calculate EX^2 and get

$$\operatorname{Var} X = e^{2\left(\mu+\sigma^2\right)} - e^{2\mu+\sigma^2}$$

Incomes are necessarily skewed to the right and modeling with a lognormal allows the use of normal-theory statistics on $\log$ (income).

Double Exponential Distribution

Reflect the exponential distribution around its mean.

$$f(x \mid \mu, \sigma) = \frac{1}{2\sigma} e^{-|x-\mu|/\sigma}, \quad -\infty < x < \infty, \quad -\infty < \mu < \infty, \quad \sigma > 0$$

The double exponential provides a symmetric distribution with "fat" tails (much fatter than the normal), but still retains all of its moments.

$$\mathrm{EX} = \mu \quad \text{and} \quad \operatorname{Var} X = 2\sigma^2$$

Location and Scale Families

$g(x \mid \mu, \sigma) = \frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right)$ is a pdf.

https://cse.buffalo.edu/~hungngo/classes/2005/Expanders/notes/asymptotic-estimates.pdf

The Poisson postulates

N_t is number of arrivals in the time period from time 0 to time t.
i. $N_0 = 0$
ii. $s < t \Rightarrow N_s$ and $N_t - N_s$ are independent
iii. N_s and $N_{t+s} - N_t$ are identically distributed
iv. $\lim_{t\to 0} \frac{P(N_t=1)}{t} = \lambda$
v. $\lim_{t\to 0} \frac{P(N_t>1)}{t} = 0$

If $i - v$ hold, then for any intiger n,

$$P\left(N_t = n\right) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}$$

that is, $N_t \sim$ Poisson (λt).