Binomial distribution *binomial*(n,p)

Let n be a positive integer and $p \in [0,1]$.
The binomial pmf with size n and probability p is

$$f(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & x = 0,1,\dots,n \\ 0 & \text{otherwise,} \end{cases}$$

- We have previously obtained that, if $X \sim f$, then
  $E(X) = np$;
  $\text{Var}(X) = np(1-p)$;
  mgf $M_X(t) = (pe^t + 1 - p)^n$, chf $\phi_X(t) = (pe^{it} + 1 - p)^n$, $t \in \mathscr{R}$.
- In the special case of $n = 1$, the binomial distribution is also called the Bernoulli distribution with probability p.
- If $X_1, \dots, X_n$ are n Bernoulli random variables from n independent Bernoulli trials with the same probability p, then $X = \sum_{i=1}^n X_i$ has the binomial distribution with size n and parameter p.

Poisson distribution $Poisson(\lambda)$

The Poisson distribution with parameter $\lambda > 0$ has pmf

$$f(x) = \begin{cases} \frac{\lambda^x e^{-\lambda}}{x!} & x = 0, 1, 2, \ldots \\ 0 & \text{otherwise}, \end{cases}$$

The mgf has been obtained previously, which is $e^{\lambda(e^t - 1)}$, $t \in \mathscr{R}$.
The chf is then $e^{\lambda(e^{it} - 1)}$, $t \in \mathscr{R}$.

If $X \sim f$, then

$$E(X) = \frac{d}{dt} e^{\lambda(e^t - 1)} \Big|_{t=0} = \lambda e^t e^{\lambda(e^t - 1)} \Big|_{t=0} = \lambda$$

$$E(X^2) = \lambda \frac{d}{dt} e^t e^{\lambda(e^t - 1)} \Big|_{t=0} = \lambda \left[e^t e^{\lambda(e^t - 1)} + e^t \lambda e^t e^{\lambda(e^t - 1)} \right] \Big|_{t=0} = \lambda + \lambda^2$$

Hence,

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = \lambda + \lambda^2 - \lambda^2 = \lambda$$

Example 3.2.4 (Waiting time)

As an example of a waiting-for-occurrence application, consider a telephone operator who, on the average, handles 5 calls per 3 minutes.

Let $X =$ number of calls in a minute.
If X follows a Poisson distribution, what is λ?
Since $E(X) =$ average of calls in a minute, $\lambda = E(X) = 5/3$.
What is the probability that there will be no calls in the next minute?

$$P(X = 0) = \frac{e^{-5/3}(5/3)^0}{0!} = e^{-5/3} = 0.189$$

What is the probability that there will be at least two calls?

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1) = 1 - 0.189 - \frac{e^{-5/3}(5/3)^1}{1!} = 0.496$$

Hypergeometric distribution *hypergeometric*(K, M, N)

Let K, M, and N be positive integers with $M \leq N$. The hypergeometric distribution has pmf

$$f(x) = \begin{cases} \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}} & x = 0, 1, ..., \min\{K, M\},\ M - N + K \leq x \\ 0 & \text{otherwise} \end{cases}$$

If we select K balls at random without replacement from a box filled with M red balls and $N - M$ green balls, then

$$f(x) = P(\text{exactly } x \text{ of the balls are red})$$

Typically, $K < \min\{M, N\}$ and then

$$f(x) = \begin{cases} \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}} & x = 0, 1, ..., K, \\ 0 & \text{otherwise} \end{cases}$$

Thus

$$\sum_{x=0}^{K} \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}} = 1 \qquad \text{(not trivial to verify)}$$

If $X \sim f$, then

$$\begin{aligned} E(X) &= \sum_{x=0}^{K} x \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}} = \sum_{x=1}^{K} \frac{KM\binom{M-1}{x-1}\binom{N-M}{K-x}}{N\binom{N-1}{K-1}} \\ &= \frac{MK}{N} \sum_{y=0}^{K-1} \frac{\binom{M-1}{y}\binom{(N-1)-(M-1)}{K-1-y}}{\binom{N-1}{K-1}} = \frac{MK}{N} \end{aligned}$$

A similar but more lengthy argument leads to

$$\mathrm{Var}(X) = \frac{KM(N-M)(N-K)}{N^2(N-1)}$$

Negative binomial *negative-binomial*(r, p)

Let r be a positive integer and $p \in [0,1]$.
The negative binomial pmf is

$$f(x) = \begin{cases} \binom{x-1}{r-1} p^r (1-p)^{x-r} & x = r, r+1, \dots \\ 0 & \text{otherwise,} \end{cases}$$

- In a sequence of independent Bernoulli trials with probability p, if X is the number of trials needed to have the rth success (1 in the Bernoulli trial), then $X \sim f$.
- If $Y = X - r$ (the number of 0's before the rth 1), then

$$P(Y = y) = \binom{r+y-1}{y} p^r (1-p)^y, \quad y = 0, 1, 2, \dots$$

  This is also called the negative binomial distribution.
- The special case of negative binomial distribution with $r = 1$ is called the geometric distribution:

$$f(x) = \begin{cases} p(1-p)^{x-1} & x = 1, 2, \dots \\ 0 & \text{otherwise,} \end{cases}$$

The expectation of the negative binomial distribution (2nd definition) is

$$\begin{aligned} E(Y) = &= \sum_{y=0}^{\infty} y \binom{r+y-1}{y} p^r (1-p)^y = \sum_{y=1}^{\infty} \frac{(r+y-1)!}{(y-1)!(r-1)!} p^r (1-p)^y \\ &= \sum_{y=1}^{\infty} r \binom{r+y-1}{y-1} p^r (1-p)^y = \sum_{z=0}^{\infty} r \binom{r+z}{z} p^r (1-p)^{z+1} \\ &= \frac{r(1-p)}{p} \sum_{z=0}^{\infty} \binom{(r+1)+(z-1)}{z} p^{r+1} (1-p)^z \\ &= \frac{r(1-p)}{p} \end{aligned}$$

The expectation of the negative binomial distribution (1st definition) is

$$E(X) = E(Y) + r = \frac{r(1-p)}{p} + r = \frac{r}{p}$$

A similar calculation shows

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

Uniform distribution $uniform(a, b)$

The uniform distribution on the interval $[a, b]$, where a and b are real numbers with $a < b$, has pdf

$$f(x) = \begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

If $X \sim f$, then

$$E(X) = \frac{1}{b-a}\int_a^b x dx = \frac{x^2}{2(b-a)}\bigg|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}$$

$$E(X^2) = \frac{1}{b-a}\int_a^b x^2 dx = \frac{x^3}{3(b-a)}\bigg|_a^b = \frac{b^3 - a^3}{2(b-a)} = \frac{b^2 + a^2 + ab}{3}$$

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = \frac{b^2 + a^2 + ab}{3} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}$$

$$M_X(t) = \frac{1}{b-a}\int_a^b e^{tx} dx = \frac{e^{tb} - e^{ta}}{t(b-a)}, \quad t \in \mathscr{R}$$

Normal distribution $N(\mu, \sigma^2)$

The normal distribution plays a central role in statistics.

- The normal distributions and distributions associated with it are very tractable analytically.
- The normal distribution has the familiar bell shape, whose symmetry makes it an appealing choice for many population models.
- Under the Central Limit Theorem (Chapter 5), the normal distribution can be used to approximate a large variety of distributions in large samples.

Let $\mu \in \mathscr{R}$ and $\sigma > 0$ be two constants.
The normal distribution $N(\mu, \sigma^2)$ has pdf

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}, \quad x \in \mathscr{R}$$

When $\mu = 0$ and $\sigma = 1$, $N(0,1)$ is called the standard normal distribution.
The following is a figure of $N(0,1)$ pdf curve.

Standard normal pdf curve

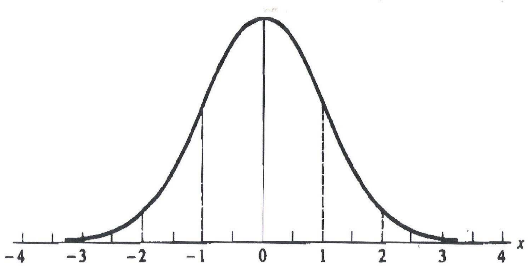


Figure 3.3.1. *Standard normal density*

We first need to show that f is indeed a pdf.

By the transformation $z = (x - \mu)/\sigma$,

$$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-(x-\mu)^2/(2\sigma^2)}dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-z^2/2}dz = \frac{2}{\sqrt{2\pi}}\int_{0}^{\infty} e^{-z^2/2}dz$$

because $e^{-z^2/2}$ is symmetric.

Note that

$$\left(\int_0^\infty e^{-z^2/2}dz\right)^2 = \left(\int_0^\infty e^{-z^2/2}dz\right)\left(\int_0^\infty e^{-u^2/2}du\right)$$
$$= \int_0^\infty \int_0^\infty e^{-(z^2+u^2)/2}dudz$$
$$= \int_0^\infty \int_0^{\pi/2} re^{-r^2/2}d\theta dr \qquad z = r\cos\theta,\ u = r\sin\theta$$
$$= \frac{\pi}{2}\int_0^\infty re^{-r^2/2}dr = \frac{\pi}{2}e^{-r^2/2}\Big|_\infty^0 = \frac{\pi}{2}$$

Thus,

$$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^\infty e^{-(x-\mu)^2/(2\sigma^2)}dx = \frac{2}{\sqrt{2\pi}}\int_0^\infty e^{-z^2/2}dz = \frac{2}{\sqrt{2\pi}}\sqrt{\frac{\pi}{2}} = 1$$

A consequence of this result is

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

Hence

$$\Gamma(1/2) = \sqrt{\pi}$$

Properties of normal distributions

- $X \sim N(\mu, \sigma^2)$ iff $Z = (X - \mu)/\sigma \sim N(0,1)$ (by transformation).
- If $Z \sim N(0,1)$, then for any $t \in \mathscr{R}$, its mgf is

$$M_Z(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tz} e^{-z^2/2} dz = \frac{e^{t^2/2}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(z-t)^2/2} dz = e^{t^2/2}$$

- If $X \sim N(\mu, \sigma^2)$, then $Z = (X - \mu)/\sigma \sim N(0,1)$ and the mgf of X is

$$M_X(t) = M_{\sigma Z + \mu}(t) = e^{\mu t} M_Z(\sigma t) = e^{\mu t + \sigma^2 t^2/2}, \quad t \in \mathscr{R}$$

- If $X \sim N(\mu, \sigma^2)$, then its chf is $\phi_X(t) = e^{\imath \mu t - \sigma^2 t^2/2}$, $t \in \mathscr{R}$.
- If $X \sim N(0, \sigma^2)$, by differentiating $M_X(t) = e^{\sigma^2 t^2/2}$, we obtain that

$$E(X^r) = \begin{cases} (r-1)(r-3)\cdots 3 \cdot 1 \sigma^r & \text{when } r \text{ is an even integer} \\ 0 & \text{when } r \text{ is an odd integer} \end{cases}$$

- If $Z \sim N(0,1)$, then $E(Z) = 0$ and $\mathrm{Var}(Z) = E(Z^2) = 1$.
- If $X \sim N(\mu, \sigma^2)$, then $E(X) = \mu$ and $\mathrm{Var}(X) = \sigma^2$.
- If $X \sim N(\mu, \sigma^2)$, the distribution of $Y = e^X$ is called the log-normal distribution with $E(Y) = e^{\mu + \sigma^2/2}$ and $\mathrm{Var}(Y) = e^{2(\mu+\sigma^2)} - e^{2\mu+\sigma^2}$.

- The cdf of the standard normal is called the standard normal cdf and denoted by

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt, \quad x \in \mathscr{R}$$

  But this cdf does not have a close form.

- If $X \sim N(\mu, \sigma^2)$, then its cdf is

$$F_X(x) = P\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = P\left(Z \le \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

  We can then use the standard normal cdf to calculate probabilities related to all normal random variables.

- The pdf of $N(\mu, \sigma^2)$ is a bell shaped curve that is symmetric about μ, maximized at μ, and changes from concave to convex at $\mu \pm \sigma$.
- When $X \sim N(\mu, \sigma^2)$,

$$\begin{aligned} P(|X-\mu| \le \sigma) &= \Phi(1) - \Phi(-1) = 0.6826 \\ P(|X-\mu| \le 2\sigma) &= \Phi(2) - \Phi(-2) = 0.9544 \\ P(|X-\mu| \le \sigma) &= \Phi(3) - \Phi(-3) = 0.9974 \end{aligned}$$

Gamma distribution $gamma(\alpha, \beta)$

Let $\alpha > 0$ and $\beta > 0$ be two constants.

The gamma distribution with shape parameter α and scale parameter β has pdf

$$f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

where

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

is called the gamma function,

$$\Gamma(1/2) = \sqrt{\pi}, \quad \Gamma(\alpha+1) = \alpha\Gamma(\alpha), \quad \Gamma(n) = (n-1)!,\ n = 1, 2, ...$$

If $\alpha < 1$, then the pdf of $gamma(\alpha, \beta)$ is decreasing in x and unbounded at $x = 0$.

If $\alpha \geq 1$, then the pdf of $gamma(\alpha, \beta)$ is bounded, increasing in $x < (\alpha-1)\beta$, and decreasing in $x > (\alpha-1)\beta$.

Previously, we showed that, if $X \sim gamma(\alpha, \beta)$, then

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

$$E(X^n) = \alpha(\alpha+1)\cdots(\alpha+n-1)\beta^n, \quad n = 1, 2, \ldots$$

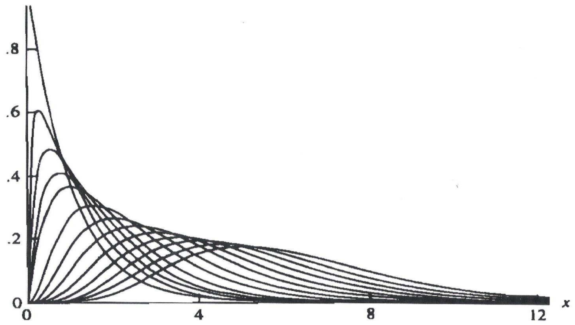
In particular,

$$E(X) = \alpha\beta, \quad E(X^2) = \alpha(\alpha+1)\beta^2, \quad \mathrm{Var}(X) = \alpha\beta^2$$

From the mfg, the chf is

$$\phi_X(t) = (1 - i\beta t)^{-\alpha}, \quad t \in \mathscr{R}$$

Some gamma pdf curves



Chi-square distribution

The gamma distribution *gamma*(α, β) with $\beta = 2$ and $\alpha = k/2$ for a positive integer k is called the chi-square distribution with k degrees of freedom, and its pdf is

$$f(x) = \begin{cases} \frac{1}{\Gamma(k/2)2^{k/2}} x^{(k/2)-1} e^{-x/2} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

- If $X \sim f$, then
$$E(X) = k, \quad \mathrm{Var}(X) = 2k,$$
and the mgf $M_X(t)$ and chf $\phi_X(t)$ are given by
$$M_X(t) = (1-2t)^{-k/2}, \ \ t < 1/2, \quad \phi_X(t) = (1-2it)^{-k/2}, \ \ t \in \mathscr{R}$$
- The chi-square distribution is closely related to the normal distribution.
- In Example 2.1.9, we showed that if $X \sim N(0,1)$, then $X^2 \sim$ chi-square with 1 degree of freedom.
- A similar result about the chi-square distribution with k degrees of freedom will be introduced later.

Exponential distribution *exponential*(β)

A special case of the gamma distribution with $\alpha = 1$ is called the exponential distribution, i.e., $gamma(1,\beta) = exponential(\beta)$.

The $exponential(\beta)$ distribution has pdf

$$f(x) = \begin{cases} \frac{1}{\beta} e^{-x/\beta} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

For any $x > 0$,

$$P(X \leq x) = \frac{1}{\beta}\int_0^x e^{-t/\beta}\,dt = e^{-t/\beta}\Big|_x^0 = 1 - e^{-x/\beta}$$

Hence, the cdf of $exponential(\beta)$ has an explicit form

$$F(x) = \begin{cases} 1 - e^{-x/\beta} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

An example of $X \sim exponential(\beta)$ is the lifetime.

Since *exponential*(β) = *gamma*$(1,\beta)$, if $X \sim$ *exponential*(β),

$$E(X) = \beta, \quad \mathrm{Var}(X) = \beta^2, \quad M_X(t) = \frac{1}{1-\beta t}, \ \ t < \frac{1}{\beta},$$

$$\phi_X(t) = \frac{1}{1-i\beta t}, \quad t \in \mathscr{R}$$

The exponential distribution has a memoryless property, i.e.,

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

for any $s > t \geq 0$, because

$$\begin{aligned} P(X > s|X > t) &= \frac{P(X > s, X > t)}{P(X > t)} = \frac{P(X > s)}{P(X > t)} \\ &= \frac{1-F(s)}{1-F(t)} = \frac{e^{-s/\beta}}{e^{-t/\beta}} = e^{-(s-t)/\beta} \\ &= 1 - F(s-t) = P(X > s-t) \end{aligned}$$

That means, if X is the lifetime of a product, then the lifetime of a product that has been used t hours has the same distribution as that of a new product (new = used).

Cauchy distribution *Cauchy*(μ,σ)

For constants $\mu \in \mathscr{R}$ and $\sigma > 0$, the *Cauchy*(μ,σ) distribution has pdf

$$f(x) = \frac{\sigma}{\pi[\sigma^2 + (x-\mu)^2]}, \quad x \in \mathscr{R}$$

The Cauchy pdf is bell-shaped and symmetric about μ.

Its biggest difference from the normal pdf is that, if $X \sim$ *Cauchy*(μ,σ), then $E|X| = \infty$ (and hence $E|X|^n = \infty$, $n = 2,3,...$)

The mgf of X is ∞ except at 0 and the chf of X is $e^{i\mu+\sigma|t|}$, $t \in \mathscr{R}$.

μ still measures the center of *Cauchy*(μ,σ) although the expectation does not exist: μ is the median in the sense that $P(X < \mu) = \frac{1}{2}$.

Standard normal and Cauchy pdf curves