Chapter 5 Distributions

Discrete Distributions

	PMF	Support	Mean	Var	MGF
Poisson(λ)	$\frac{e^{-\lambda}\lambda^K}{K!}$	$0,1,2,\ldots$	λ	λ	$e^{\lambda(e^t-1)}$
Binomial(n,p)	$\binom{n}{K}p^K(1-p)^{n-K}$	$0,1,\ldots,n$	np	$np(1-p)$	$(1-p+pe^t)^n$
Geometric(p)	$p(1-p)^K$	$0,1,2,\ldots$	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$	$\frac{p}{1-(1-p)e^t}$
Neg Bin(r,p)	$\binom{r+K-1}{K}p^r(1-p)^K$	$0,1,2,\ldots$	$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$	$\left[\frac{p}{1-(1-p)e^t}\right]^r$
Hypergeo(A,B,n)	$\frac{\binom{A}{K}\binom{B}{n-K}}{\binom{A+B}{n}}$	$\max(0,n-B)\le K\le \min(n,A)$	$\frac{nA}{A+B}$	$\frac{nAB}{(A+B)^2}\left(\frac{A+B-n}{A+B-1}\right)$	—

Let's Pick off some details

$X \sim Bin(n,p)$ MGF

$X = X_1 + X_2 + \ldots + X_n$

$$X_i = \begin{cases} 1, & \text{trial } i \text{ a success} \\ 0, & \text{trial } i \text{ a failure} \end{cases}$$

$$\psi_X(t) = E\left[e^{tX}\right] = E\left[e^{t(X_1+\ldots+X_n)}\right]$$

$$= E\left[e^{tX_1}\right]E\left[e^{tX_2}\right]\ldots E\left[e^{tX_n}\right]$$

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

$$E\left[e^{tX_i}\right] = e^{t\cdot 0}P(X_i=0) + e^{t\cdot 1}P(X_i=1) = 1-p+pe^t$$

$$\psi_X'(t) = n(1-p+pe^t)^{n-1}pe^t$$

$$\psi_x'(0) = n(1-p+pe^0)^{n-1} pe^0 = np = E[X]$$

$$\psi_x''(t) = n(1-p+pe^t)^{n-1} pe^t + pe^t n(n-1)(1-p+pe^t)^{n-2} pe^t$$

$$\psi_x''(0) = np + p^2 n(n-1)$$

$$Var(X) = \psi_x''(0) - (\psi_x'(0))^2 = np + n^2p^2 - np^2 - (np)^2$$
$$= np(1-p)$$

■

$X \sim Geo(p)$ $\quad p \in (0,1)$ $\quad$ (Book Def)

$$\psi_X(t) = E[e^{tX}] = \sum_{K=0}^{\infty} e^{tK} p(1-p)^K$$
$$= p \sum_{K=0}^{\infty} [(1-p)e^t]^K$$
$$= p \frac{1}{1-(1-p)e^t} \quad \text{if } (1-p)e^t < 1 \iff t < \log\left(\frac{1}{1-p}\right)$$
$$= \frac{p}{1-qe^t}$$

This is not the same geometric distribution that I gave you earlier that was supported in $\{1, 2, \dots\}$

Often, this distribution is called the "Zero-modified" geometric distribution.

$$\psi_x'(t) = p(-1)(1-qe^t)^{-2}(-qe^t)$$

$$E[X] = \psi_X'(0) = pq/(1-q)^2 = \frac{p(1-p)}{p^2} = \frac{1-p}{p}$$

For You: get $Var(X) = \frac{1-p}{p^2}$ through the MGF

Interpretation: $X \sim MGeo(p)$ counts the # of failures in independent trials until the first success

<u>Fact</u> The Geo RV is the only memoryless discrete RV supported on $\{1, 2, \ldots\}$

Memoryless $\quad P(X > k+n \mid X > n) = P(X > k)$

<u>Why</u>

X Memoryless $\Rightarrow$

$$P(X > k+n \mid X > n) = P(X > k) \iff$$

$$\frac{P(X > k+n \cap X > n)}{P(X > n)} = P(X > k) \iff$$

$$P(X > k+n) = P(X > k)\, P(X > n) \iff$$

$$\eta_{k+n} = \eta_k \eta_n \qquad \eta_k = P(X > k) \iff$$

$$a_{k+n} = a_k + a_n \qquad a_k = \log(P(X > k))$$

$$a_1 = \log(P(X > 1))$$

$$a_2 = a_1 + a_1 = 2a_1, \ldots, a_k = k a_1$$

So $\log(P(X>K)) = K a_1 \Rightarrow$

$$P(X>K) = \left(e^{a_1}\right)^K$$

Now $P(X=K) = P(X>K-1) - P(X>K)$

$$= \left(e^{a_1}\right)^{K-1} - \left(e^{a_1}\right)^K = \left(e^{a_1}\right)^{K-1}\left[1 - e^{a_1}\right]$$

Has form pq^{K-1}, where $p = 1 - e^{a_1} = 1 - e^{\log(P(X>1))}$

$$= 1 - P(X>1) = P(X=1)$$

$\boxtimes$

Neg Bin(r, p) $\quad r \in \{1, 2, \ldots\}$

$$P(X=K) = \binom{r+K-1}{K} p^r (1-p)^K, \quad K = 0, 1, 2, \ldots$$

This represents the # of failures in indep coin flips until we get r successes

	1	2	3	4	5	6	
$r=2$	T	H	T	T	T	H	$X=4$

$P(X=K) = P($ K failures in first $r+K-1$ trials $\cap$ Trial $r+K$ is a success $)$

$$= \binom{r+K-1}{K} (1-p)^K p^{r-1} \cdot p$$

$$= \binom{r+K-1}{K} p^r (1-p)^K$$

Write $X = X_1 + X_2 + \dots + X_r$

X_i = # failures until i^{th} success $\sim$ M Geo(p)

So $\psi_X(t) = \psi_{X_1}(t) \dots \psi_{X_r}(t) = \left[\dfrac{p}{1-(1-p)e^t}\right]^r$

It follows that $E[X] = \dfrac{r(1-p)}{p}$ & $Var(X) = \dfrac{r(1-p)}{p^2}$

□

Hypergeometric Distributions are the messiest!

$X \sim$ Hyper Geo (A, B, n)

There are A Black balls (Type I) in a box
B White balls (Type II)

n are drawn out at random <u>without replacement</u>

X is the # of Black balls drawn:

$$P(X=K) \quad \frac{\binom{A}{K}\binom{B}{n-K}}{\binom{A+B}{n}}$$

Constraints on the possible X values

black balls drawn $\geq n-B$; # black balls drawn $\leq A$

$$\max(0, n-B) \leq K \leq \min(n, A)$$

Let $X_i = \begin{cases} 1, & \text{draw } i \text{ is black} \\ 0, & \text{draw } i \text{ is white} \end{cases}$

$$X = X_1 + \dots + X_n$$

$$P(X_1 = 1) = \frac{A}{A+B}$$

$$P(X_2 = 1) = \frac{A}{A+B}$$

(we made a big deal about this earlier)

$$\vdots$$

$$P(X_n = 1) = \frac{A}{A+B}$$

$$E[X_i] = 0\,P(X_i = 0) + 1\,P(X_i = 1) = \frac{A}{A+B}$$

So $E[X] = E[X_1] + \dots + E[X_n] = n\left(\frac{A}{A+B}\right)$

$$\begin{aligned} Var(X) &= Var(X_1 + \dots + X_n) \\ &= \sum_{i=1}^{n} Var(X_i) + 2 \sum\sum_{1 \le i < j \le n} Cov(X_i, X_j) \end{aligned}$$

$$\begin{aligned} Var(X_i) &= E[X_i^2] - E[X_i]^2 = E[X_i] - E[X_i]^2 \\ &= \frac{A}{A+B}\left(1 - \frac{A}{A+B}\right) = \frac{A}{A+B}\,\frac{B}{A+B} \end{aligned}$$

$i < j$

$$Cov(X_i, X_j) = E[X_i X_j] - E[X_i]E[X_j]$$

$$\begin{aligned} E[X_i X_j] = P(X_i X_j = 1) &= P(X_j = 1 \mid X_i = 1)\,P(X_i = 1) \qquad j > i \\ &= \frac{A-1}{A-1+B} \quad \frac{A}{A+B} \qquad \text{why} \end{aligned}$$

So $Cov(X_i, X_j) = \frac{A}{A+B}\,\frac{A-1}{A-1+B} - \left(\frac{A}{A+B}\right)^2 =$

$$\frac{A}{A+B}\left[\frac{A-1}{A-1+B} - \frac{A}{A+B}\right] = \frac{A}{A+B}\left[\frac{(A+B)(A-1) - A(A+B-1)}{(A+B)(A-1+B)}\right]$$

$$= -\frac{AB}{(A+B)^2(A+B-1)}$$ Note negative sign

So $$Var(X) = n\frac{AB}{(A+B)^2} - \frac{n(n-1)AB}{(A+B)^2(A+B-1)}$$

$$= \frac{nAB}{(A+B)^2}\left(\frac{A+B-n}{A+B-1}\right)$$

$$= \text{"}npq\text{"}\underbrace{\left(\frac{A+B-n}{A+B-1}\right)}_{\text{dependence factor for the dependent trials}}$$ $p = \frac{A}{A+B}$ $q = 1-p$

MGF : No known closed form.

Poisson vs. Binomial

Suppose $X_n \sim \text{Binomial}(n, p_n)$. Then if $np_n \to \lambda > 0$,

$$\lim_{n\to\infty} P[X_n = k] = \frac{e^{-\lambda}\lambda^k}{k!}$$

Why

$$\lim_{n\to\infty} P[X_n = k] = \lim_{n\to\infty} \binom{n}{k} p_n^k (1-p_n)^{n-k}$$

$$= \lim_{n\to\infty} \frac{n!}{k!(n-k)!} p_n^k (1-p_n)^{n-k}$$

$$= \lim_{n\to\infty} \frac{n(n-1)\dots(n-k+1)}{k!}\left(\frac{\lambda_n}{n}\right)^k\left(1-\frac{\lambda_n}{n}\right)^{n-k}$$

Let $\lambda_n = np_n$

$\lambda_n \to \lambda$

$$= \lim_{n\to\infty} \frac{n}{n}\frac{n-1}{n}\cdots\frac{n-k+1}{n} \frac{\lambda_n^k\left(1-\frac{\lambda_n}{n}\right)^n\left(1-\frac{\lambda_n}{n}\right)^{-k}}{k!}$$

Now $\frac{n-j}{n} \to 1$ for each $j \in \{1, 2, \ldots, k-1\}$ as $n \to \infty$

$$\left(1 - \frac{\lambda_n}{n}\right)^{-k} = \left(1 - p_n\right)^{-k} \to 1 \quad \text{as} \quad n \to \infty$$

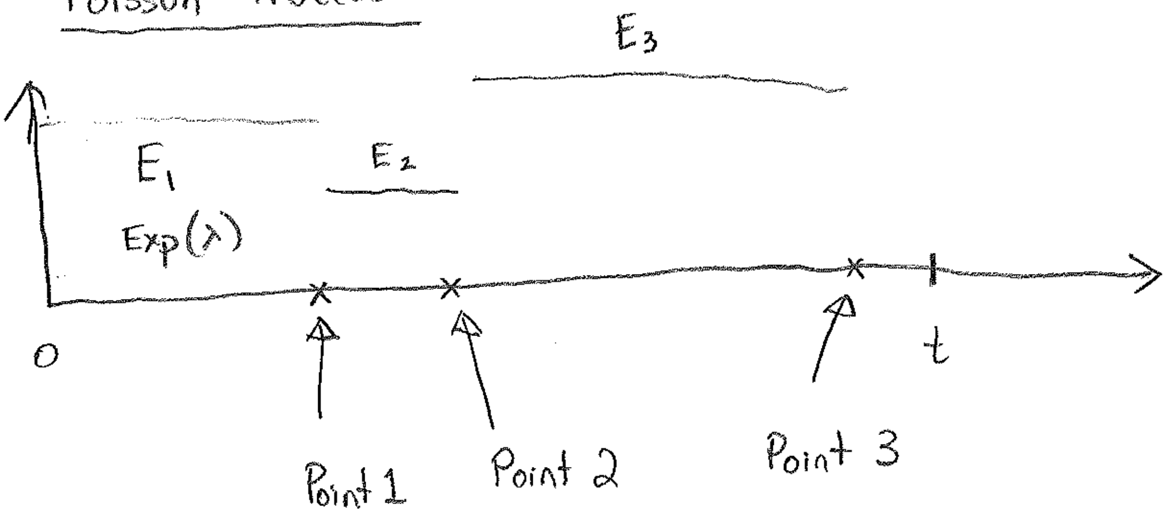
$$\lambda_n^{\,k} \to \lambda^k \quad \text{as} \quad n \to \infty$$

$$\left(1 - \frac{\lambda_n}{n}\right)^{n} \to e^{-\lambda} \qquad \left(1 - \frac{x}{n}\right)^{n} \to e^{-x}$$

(See Theorem 5.3.3 in text)

So $$\lim_{n \to \infty} P[X_n = k] = \frac{e^{-\lambda} \lambda^k}{k!}, \qquad k = 0, 1, 2, \ldots$$

<u>Poisson Process</u>



Assumption $\{E_i\}_{i=1}^{\infty}$ are IID $Exp(\lambda)$

Let $N(t)$ be the # of points that occur in $(0, t]$

Fact: $N(t) \sim \text{Poisson}(\lambda t)$ Maybe proof later

Continuous Distributions

$X \sim \text{Gamma}(\alpha, \beta)$ $\quad f_X(x) = \dfrac{x^{\alpha-1} e^{-\beta x} \beta^{\alpha}}{\Gamma(\alpha)}, \quad x > 0$

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

Facts

1) $\Gamma(\alpha) < \infty$ if $\alpha > 0$
2) $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$
3) $\Gamma(n) = (n-1)!$, n an integer

Why

1) For You
2) Integration by parts gives

$$\int_0^\infty x^{\alpha-1} e^{-x}\, dx = \qquad \begin{array}{cc} x^{\alpha-1} & e^{-x} \\ (\alpha-1)x^{\alpha-2} & -e^{-x} \end{array}$$

$$-e^{-x} x^{\alpha-1} \Big|_{x=0}^{x=\infty} + \int_0^\infty (\alpha-1) x^{\alpha-2} e^{-x}\, dx$$

$$= 0 + (\alpha-1)\Gamma(\alpha-1).$$

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

$\Gamma(2) = (2-1)\Gamma(2-1) = 1\,\Gamma(1) = 1$

$\Gamma(3) = (3-1)\Gamma(3-1) = 2\,\Gamma(2) = 2!$

$\vdots$ Induct

Going back to the Gamma distribution, note that

$$\int_0^\infty \frac{x^{\alpha-1} e^{-\beta x} \beta^\alpha}{\Gamma(\alpha)} dx = \int_0^\infty \frac{\left(\frac{y}{\beta}\right)^{\alpha-1} e^{-y} \beta^\alpha}{\Gamma(\alpha)} \frac{dy}{\beta}$$

$y = \beta x$
$dy = \beta dx$

$$= \frac{1}{\Gamma(\alpha)} \int_0^\infty y^{\alpha-1} e^{-y} dy = \frac{\Gamma(\alpha)}{\Gamma(\alpha)} = 1$$

MGF:

$$E\left[e^{tX}\right] = \int_0^\infty e^{tx} \frac{x^{\alpha-1} e^{-\beta x} \beta^\alpha}{\Gamma(\alpha)} dx$$

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

let $y = (\beta - t)x$
$dy = (\beta - t)dx$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty \left(\frac{y}{\beta-t}\right)^{\alpha-1} e^{-y} \frac{dy}{\beta-t}$$

$$= \left(\frac{\beta}{\beta-t}\right)^\alpha \frac{1}{\Gamma(\alpha)} \int_0^\infty y^{\alpha-1} e^{-y} dy = \left(\frac{\beta}{\beta-t}\right)^\alpha$$

Only finite when $(\beta - t) > 0 \Rightarrow t < \beta$.

$$\psi_X'(t) = \alpha \left(\frac{\beta}{\beta-t}\right)^{\alpha-1} \beta(-1)(\beta-t)^{-2}(-1)$$

$$\psi_X'(0) = \alpha\left(\frac{\beta}{\beta-0}\right)^{\alpha-1} \frac{\beta}{(\beta-0)^2} = \frac{\alpha}{\beta} = E[X]$$

For You: Get $Var(X) = \frac{\alpha}{\beta^2}$

The $Exp(\beta)$ distribution results when $\alpha = 1$

$$f_X(x) = \beta e^{-\beta x}, \quad x > 0$$

The $Exp(\beta)$ is the only cts distribution supported on $[0,\infty)$

Why

If $P(X > x+y \mid X > y) = P(X > x) \quad \forall\, x, y \geq 0$

$\Rightarrow \quad P(X > x+y) = P(X > x)\, P(X > y)$

let $\bar{F}(x) = 1 - F_X(x)$

$$\ln\left(\bar{F}(x+y)\right) = \ln\left(\bar{F}(x)\right) + \ln\left(\bar{F}(y)\right)$$

Has form $\quad \eta(x+y) = \eta(x) + \eta(y)$

The only cts solution to this is $\eta(x) = cx$ for some constant c.

So $\quad \ln\left(\bar{F}(x)\right) = cx$

$\Rightarrow \quad \bar{F}(x) = e^{cx}$

$\Rightarrow \quad F_X(x) = 1 - e^{cx} \quad$ or $\quad X \sim Exp(-c)$

Fact $\quad X_1 \sim Gamma(\alpha_1, \beta),\ X_2 \sim Gamma(\alpha_2, \beta), \ldots, X_n \sim Gamma(\alpha_n, \beta)$ Independent

$\Rightarrow \quad X_1 + X_2 + \cdots + X_n \sim Gamma(\alpha_1 + \cdots + \alpha_n, \beta)$

Why

$$\psi_{X_1+X_2+\cdots+X_n}(t) = E\left[e^{t(X_1+\cdots+X_n)}\right]$$

$$= \varphi_{X_1}(t) \cdots \varphi_{X_n}(t) =$$

$$\left(\frac{\beta}{\beta-t}\right)^{\alpha_1}\left(\frac{\beta}{\beta-t}\right)^{\alpha_2}\cdots\left(\frac{\beta}{\beta-t}\right)^{\alpha_n} =$$

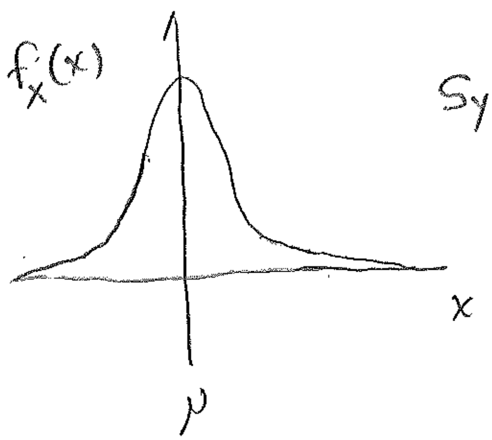
$$\left(\frac{\beta}{\beta-t}\right)^{\alpha_1+\alpha_2+\cdots+\alpha_n}$$

This is the Gamma $(\alpha_1+\alpha_2+\cdots+\alpha_n, \beta)$ MGF.

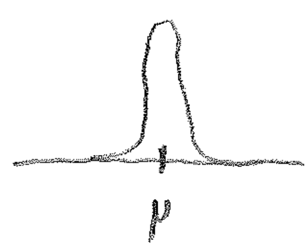
The Normal (Univariate) Distribution

$X \sim N(\mu, \sigma^2)$

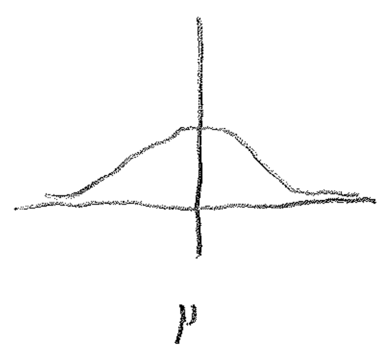
$$f_X(x) = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}, \quad -\infty < x < +\infty$$



Symmetric about μ



σ small



σ bigger

To see that $\int_{-\infty}^{+\infty} f_X(x)\,dx = 1$, note that

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

$$z = \frac{x-\mu}{\sigma}$$

To see that

$$I := \int_{-\infty}^{+\infty} \frac{e^{-z^2/2}}{\sqrt{2\pi}}\, dz \,, \qquad \text{note that}$$

$$I^2 = \int_{-\infty}^{+\infty} \frac{e^{-x^2/2}}{\sqrt{2\pi}}\, dx \int_{-\infty}^{+\infty} \frac{e^{-y^2/2}}{\sqrt{2\pi}}\, dy$$

$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-(x^2+y^2)/2}\, dy\, dx$$

$$= \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty} e^{-r^2/2}\, r\, dr\, d\theta \qquad \text{let } r = \sqrt{x^2+y^2}, \quad dy\,dx = r\,dr\,d\theta$$

$$= \frac{1}{2\pi}\int_0^{2\pi}\left[-e^{-r^2/2}\Big|_{r=0}^{r=\infty}\right] d\theta$$

$$= \frac{1}{2\pi}\int_0^{2\pi} 1\, d\theta = \underline{1} \qquad \blacksquare$$

<u>Thm</u> If $X \sim N(\mu, \sigma^2)$, $Z = \frac{X-\mu}{\sigma} \sim N(0,1) \longleftarrow$ Standard Normal

<u>Why</u>

$$F_Z(z) = P[Z \le z] = P\left[\frac{X-\mu}{\sigma} \le z\right] = P[X \le \mu + \sigma z]$$

$$= \int_{-\infty}^{\mu+\sigma z} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \qquad \text{let } y = \frac{x-\mu}{\sigma}$$

$$= \int_{-\infty}^{z} \frac{e^{-y^2/2}}{\sqrt{2\pi}}\, dy$$

Standard Normal PDF

$$f_Z(z) = \frac{e^{-z^2/2}}{\sqrt{2\pi}}$$

We give the special symbols

$$\phi(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}$$

$$\Phi(x) = \int_{-\infty}^{x} \frac{e^{-t^2/2}}{\sqrt{2\pi}}$$ for the Standard normal CDF & PDF

$$E[Z] = \int_{-\infty}^{+\infty} z \frac{e^{-z^2/2}}{\sqrt{2\pi}} \, dz = 0$$

$$Var(Z) = \int_{-\infty}^{+\infty} z^2 \frac{e^{-z^2/2}}{\sqrt{2\pi}} \, dz =$$

$$\begin{array}{cc} z & \dfrac{z e^{-z^2/2}}{\sqrt{2\pi}} \\ 1 & \dfrac{-e^{-z^2/2}}{\sqrt{2\pi}} \end{array}$$

$$z \frac{e^{-z^2/2}}{\sqrt{2\pi}} \Bigg|_{z=-\infty}^{z=+\infty} + \int_{-\infty}^{+\infty} \frac{e^{-z^2/2}}{\sqrt{2\pi}} \, dz$$

$$= 0 + 1 = 1$$

So $E\left[\frac{X-\mu}{\sigma}\right] = 0 \Rightarrow E[X-\mu] = 0 \Rightarrow E[X] = \mu$

$Var\left(\frac{X-\mu}{\sigma}\right) = 1 \Rightarrow \frac{1}{\sigma^2} = Var(X) \Rightarrow Var(X) = \sigma^2$

MGFs

$$E\left[e^{tZ}\right] = \int_{-\infty}^{+\infty} e^{tz} \frac{e^{-z^2/2}}{\sqrt{2\pi}} \, dz$$

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

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

$$= e^{t^2/2} \underbrace{\int_{-\infty}^{+\infty} \frac{e^{-\frac{1}{2}\left(\frac{z-t}{1}\right)^2}}{1\sqrt{2\pi}}}_{N(t,1) \text{ PDF}} dz = e^{t^2/2}$$

$$\psi_Z(t) = e^{t^2/2}$$

$X \sim N(\mu, \sigma^2)$ $\quad Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

$$\psi_X(t) = E[e^{tX}]$$

$$= E[e^{t(\sigma Z + \mu)}] = e^{t\mu} E[e^{t\sigma Z}]$$

$$= e^{t\mu} \psi_Z(\sigma t)$$

$$\psi_X(t) = e^{t\mu} e^{\sigma^2 t^2/2}$$

$$= \exp\{t\mu + \sigma^2 t^2/2\}$$

If $X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$, $\ldots$, $X_n \sim N(\mu_n, \sigma_n^2)$

Independent

$$X_1 + X_2 + \ldots + X_n \sim N(\mu_1 + \ldots + \mu_n, \sigma_1^2 + \ldots + \sigma_n^2)$$

<u>Why</u>

$$\psi_{X_1 + \ldots + X_n}(t) = E[e^{t(X_1 + \ldots + X_n)}] = \psi_{X_1}(t) \ldots \psi_{X_n}(t)$$

$$= \prod_{i=1}^{n} e^{t\mu_i} e^{t^2 \sigma_i^2/2}$$

$$= e^{t(\mu_1+\dots+\mu_n) + \frac{t^2}{2}(\sigma_1^2+\dots+\sigma_n^2)}$$

Which is the Normal $(\mu_1+\dots+\mu_n,\ \sigma_1^2+\dots+\sigma_n^2)$ MGF $\square$

If $X \sim N(\mu, \sigma^2)$, $Y = aX+b$ is also normal:

$$\psi_Y(t) = E\left[e^{tY}\right] = E\left[e^{t(aX+b)}\right] = e^{tb}\,\psi_X(at)$$

$$= e^{tb} \exp\left\{ at\mu + \frac{\sigma^2}{2}(at)^2 \right\}$$

$$= e^{t(b+a\mu) + \frac{t^2}{2}\sigma^2 a^2}$$

$N(b+a\mu,\ \sigma^2 a^2)$ MGF. ///

$X \sim$ log-normal (μ, σ^2) means $X = e^Y$, $Y \sim N(\mu, \sigma^2)$

The log of X is normally distributed

$\alpha, \beta > 0$

$X \sim \text{Beta}(\alpha, \beta)$ $\qquad f_X(x) = \dfrac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}, \quad 0 \le x \le 1$

$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx$ is called the Beta function

Fact: $B(\alpha,\beta) = \dfrac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ No proof now

$$E[X] = \int_{-\infty}^{+\infty} x f_X(x)\,dx = \int_0^1 x \,\frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}\,dx$$

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

$$= \frac{\Gamma(\alpha+1)\,\cancel{\Gamma(\beta)}}{\Gamma(\alpha+1+\beta)} \Big/ \frac{\Gamma(\alpha)\,\cancel{\Gamma(\beta)}}{\Gamma(\alpha+\beta)}$$

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

Likewise, $E[X^2] = \int_0^1 x^2 \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}\,dx$

$$= B(\alpha+2,\beta) / B(\alpha,\beta)$$

$$= \frac{(\alpha+1)\Gamma(\alpha+1)\,\Gamma(\beta)}{\Gamma(\alpha+2+\beta)} \Big/ \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

$$= \frac{\alpha(\alpha+1)\,\Gamma(\alpha)\Gamma(\beta)}{(\alpha+1+\beta)(\alpha+\beta)\,\Gamma(\alpha+\beta)} \Big/ \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

$$= \frac{\alpha(\alpha+1)}{(\alpha+\beta+1)(\alpha+\beta)}$$

So $Var(X) = \frac{\alpha(\alpha+1)}{(\alpha+\beta+1)(\alpha+\beta)} - \left(\frac{\alpha}{\alpha+\beta}\right)^2$

$$= \frac{(\alpha^2+\alpha)(\alpha+\beta) - \alpha^2(\alpha+\beta+1)}{(\alpha+\beta+1)(\alpha+\beta)^2}$$

$$= \frac{\cancel{\alpha^3}+\cancel{\alpha^2}+\cancel{\alpha^2\beta}+\alpha\beta-\cancel{\alpha^3}-\cancel{\alpha^2\beta}-\cancel{\alpha^2}}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

$$= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

The MGF

$$\psi_X(t) = \int_0^1 e^{tx} \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}\, dx$$

has no Known closed form.

Stray tidbit $\Gamma(\tfrac{1}{2}) = \sqrt{\pi}$

Why

$$\Gamma(\tfrac{1}{2}) = \int_0^\infty x^{\frac{1}{2}-1} e^{-x}\, dx \qquad \text{let } x = y^2/2$$

$$dx = y\,dy$$

$$= \int_0^\infty \frac{\sqrt{2}}{y} e^{-y^2/2}\, \cancel{y}\, dy \qquad \sqrt{\frac{1}{x}} = \sqrt{\frac{2}{y^2}}$$

$$= \sqrt{2}\sqrt{2\pi} \int_0^\infty \frac{e^{-y^2/2}}{\sqrt{2\pi}}\, dy$$

$$= \sqrt{2}\sqrt{2\pi}\left(\frac{1}{2}\right)$$

$$= \sqrt{\pi} \qquad \blacksquare$$

This can be useful at times

Multivariate Normal

$$\vec{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} \sim N_n(\vec{\mu}, \Sigma) \qquad \vec{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} \qquad \Sigma = (\Sigma_{ij})_{ij=1}^{n}$$

$$\Sigma_{ij} = \Sigma_{ji}$$

$$f_{\vec{X}}(\vec{x}) = (2\pi)^{-n/2} \det(\Sigma)^{-1/2} \exp\left\{-\tfrac{1}{2}(\vec{x}-\vec{\mu})^T \Sigma^{-1} (\vec{x}-\vec{\mu})\right\}$$

Some facts

1. This is a legitimate density. (maybe proof later)

2. The matrix $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1n} \\ \Sigma_{21} & \Sigma_{22} & \cdots & \Sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{n1} & \Sigma_{n,2} & \cdots & \Sigma_{nn} \end{pmatrix}$

is called a covariance matrix

Thm from linear algebra

$$P \Sigma P^T = D$$

P is a matrix such that $PP^T = P^TP = I$

D is a diagonal matrix :

$$D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

P has rows/columns that contain the standardized eigenvalues of Σ

D contains the eigenvalues of Σ

Claim All eigenvalues of a covariance matrix are ≥ 0

Now $P\Sigma P^T = D \Rightarrow P^TP\Sigma P^TP = P^TDP$

$$\Rightarrow \Sigma = P^TDP$$

$$= P^TD^{1/2}D^{1/2}P$$

$$= \underbrace{P^TD^{1/2}P}_{A}\,\underbrace{P^TD^{1/2}P}_{A^T}$$

$$D^{\frac{1}{2}} = \begin{bmatrix} \sqrt{\lambda_1} & & & 0 \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ 0 & & & \sqrt{\lambda_n} \end{bmatrix}$$

So $\Sigma = AA^T$

Moments

$E[X_i] = \mu_i$; $Cov(X_i, X_j) = \Sigma_{i,j}$

So $Var(X_i) = \Sigma_{i,i}$ & variances are on the main diagnol of the covariance matrix

No proof now

If λ & $\vec{x}$ is an eigenvalue and eigenvectors of Σ

$$\Sigma \vec{x} = \lambda \vec{x} \Rightarrow$$

$$\vec{x}^T \Sigma \vec{x} = \vec{x}^T \lambda \vec{x} = \lambda \|\vec{x}\|^2$$

So $\lambda = \dfrac{\vec{x}^T \Sigma \vec{x}}{\|\vec{x}\|^2}$

But $0 \leq Var(\vec{x}^T \vec{X}) = \vec{x}^T \Sigma \vec{x}$

Implies that $\lambda \geq 0$

$$\det(\Sigma) = \det(\Sigma P P^T) = \det(P^T \Sigma P) = \det(D)$$
$$= \prod_{i=1}^{n} \lambda_i$$

So Σ^{-1} exists if $\lambda_i > 0 \ \forall i$.

If Σ^{-1} does not exist, one of the X_is is a linear combination of the other X_is...

Say $X_1 = \alpha_2 X_2 + \cdots + \alpha_n X_n + C_1$ for some constants $\alpha_2, \alpha_3, \ldots, \alpha_n$ & C_1 (No proof)

$\Sigma^{-1/2}$ refers to a symmetric matrix with

$$\Sigma^{-1/2}\Sigma^{-1/2} = \Sigma^{-1}$$

Fact If $\vec{X} \sim N_n(\vec{\mu}, \Sigma)$ and A is an $m \times n$ matrix and $\vec{b}$ an $m \times 1$ vector,

$$A\vec{X} + \vec{b} \sim N_n(A\vec{\mu} + \vec{b},\ A\Sigma A^T)$$

So $\vec{X} \sim N_n(\vec{\mu}, \Sigma) \Rightarrow \vec{Z} = \Sigma^{-1/2}(\vec{X} - \vec{\mu})$ has

$$\vec{Z} \sim N_n\left(\vec{0},\ \Sigma^{-1/2}\Sigma\,\Sigma^{-1/2}\right) \overset{\mathcal{D}}{=} N_n(0, I_n)$$

That is, $\vec{Z} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$ has independent standard normal components

$$f_{\vec{Z}}(\vec{z}) = (2\pi)^{-1/2} e^{-z_1^2/2} \times \cdots \times (2\pi)^{-1/2} e^{-z_n^2/2}$$

$$= (2\pi)^{-n/2} \exp\left\{-\tfrac{1}{2}\sum_{i=1}^{n} z_i^2\right\}$$

MGFs

Joint MGF $\qquad \vec{X} \sim N_n(\vec{\mu}, \Sigma)$

$$\psi_{X_1,\ldots,X_n}(t_1,\ldots,t_n) = E\left[e^{t_1X_1 + \cdots + t_nX_n}\right]$$

$$= e^{\vec{t}^T\vec{\mu} + \frac{1}{2}\vec{t}^T\Sigma\vec{t}}$$

No proof

$$\vec{t} = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix}$$

What good a joint MGF is may not be clear

Thm Conditional distributions of joint normal RVs are again joint normal.

$$\vec{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_m \\ \hline X_{m+1} \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} \vec{X}^{(1)} \\ \hline \vec{X}^{(2)} \end{pmatrix} \sim N(\vec{\mu}, \Sigma)$$

Write
$$\Sigma = \left(\begin{array}{ccc|ccc} \Sigma_{11} & \cdots & \Sigma_{1m} & \Sigma_{1\,m+1} & \cdots & \Sigma_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \Sigma_{m1} & \cdots & \Sigma_{mm} & \Sigma_{m\,m+1} & \cdots & \Sigma_{mn} \\ \hline \Sigma_{m+1\,1} & & \Sigma_{m+1\,m} & \Sigma_{m+1\,m+1} & \cdots & \Sigma_{m+1\,n} \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ \Sigma_{n1} & \cdots & \Sigma_{nm} & \Sigma_{n\,m+1} & \cdots & \Sigma_{nn} \end{array}\right) = \left(\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right)$$

$$\vec{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \\ \hline \mu_{m+1} \\ \vdots \\ \mu_n \end{pmatrix} = \begin{pmatrix} \vec{\mu}^{(1)} \\ \hline \vec{\mu}^{(2)} \end{pmatrix}$$

Then
$$\vec{X}^{(1)} \mid \vec{X}^{(2)} \sim N_m\left(\vec{\mu}^{(1)} + \Sigma_{12}\Sigma_{22}^{-1}(\vec{X}^{(2)} - \vec{\mu}^{(2)}),\ \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)$$

if Σ_{22}^{-1} exists

Note: $\Sigma_{12}^{T} = \Sigma_{21}$

$\mathrm{Var}(\vec{X}^{(1)} \mid \vec{X}^{(2)}) = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ does not depend on $\vec{X}^{(2)}$

Thm X_i & X_j are independent iff $\Sigma_{ij} = 0$