What to study today (Oct 7, 2020)?

2. Probability and Distribution (Chp 1-3)

2.1 Probability (Chp1.1-4)

2.2 Random Variable and Distribution (Chp1.5-10)

2.3 Multivariate Distribution (Chp2)

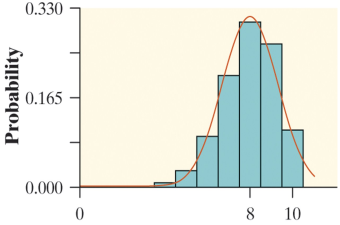
2.4 Some Important Distributions (Chp3)

- *2.4.1 Discrete Distributions*
- **2.4.2 Continuous Distributions**
- **2.4.3 Multivariate Distributions**
- **2.4.4 Distributions Induced from Others**

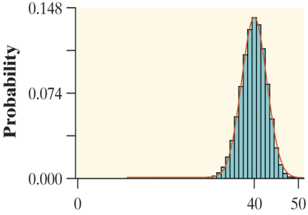
Recall that, if $X \sim B(n, \theta)$, its pmf is

$P(X = x) = \begin{pmatrix} n \\ x \end{pmatrix} \theta^x (1 - \theta)^{n-x}, \quad x = 0, \ldots, n$. When $n >> 1$, it is hard to calculate associated quantities in general.

As n gets larger, something interesting happens to the shape of a binomial distribution $B(n, \theta)$:



n = 10, *p* = 0.8



n = 50, *p* = 0.8

Proposition. Consider r.v. $X \sim B(n, p)$ where $np \geq 5$ and $n(1 - p) \geq 5$. Then $X \sim N(\mu, \sigma^2)$ approximately with $\mu = np, \sigma^2 = np(1 - p)$.

2.4.2 Continuous Distributions: Exponential Distribution

Definition. A r.v. X has an **exponential** distribution with $\lambda > 0$, denoted by $X \sim \mathit{Exponential}(\lambda)$ or $NE(\lambda)$ if it has pdf

$$f(x) = \lambda e^{-\lambda x}, \quad x > 0.$$

- The pdf is decreasing for $x > 0$, and asymmetric.
- The cdf is $F(x) = 1 - e^{-\lambda x}$ for $x > 0$.
- $E(X) = 1/\lambda$ and $V(X) = 1/\lambda^2$.
- $NE(\lambda)$ is a special case, when $\alpha = 1, \beta = 1/\lambda$, of the Gamma distribution $\mathit{Gamma}(\alpha, \beta)$:

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

$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$ and $\alpha > 0, \beta > 0$.

More on the exponential distn ... ...

- The exponential distribution has the *memoryless* property:

$$P(X > a + b \big| X > a) = P(X > b), \quad a > 0, \; b > 0.$$

e.g. Suppose that the lifespan of a lightbulb in hours $X \sim NE(\lambda)$. The prob of a used lightbulb (that has already lasted a hours) lasts an additional b hours or more is the same as a new lightbulb does.

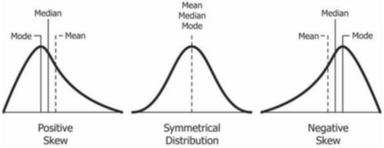
- Recall $\{N(t), t > 0\}$ is a Poisson process with the rate of λ, (the number of events over $[0, 1]$) $X = N(1) \sim Poisson(\lambda)$. In fact, $N(t) \sim Poisson(\lambda t)$.
  Y=the waiting time until the first event follows $NE(\lambda)$:

$$P(Y \leq y) = 1 - P\big(N(y) = 0\big) = 1 - e^{-\lambda y}.$$

More on Univariate Distributions

Consider rv X,

- **expectation** (the population mean) $E(X) = \int x dF(x)$, often denoted by μ
  - **median** m satisfies $P(X \leq m) \geq 1/2$ and $P(X \geq m) \geq 1/2$
- **variance** $Var(X) = E\left[(X-\mu)^2\right] = E(X^2) - (EX)^2$, often denoted by σ^2; **standard deviation** $\sigma = \sqrt{Var(X)}$.
- **skewness** $E\left\{\left[\frac{X-\mu}{\sigma}\right]^3\right\}$: *left-tailed, symmetric, right-tailed*



– *Wikipedia*

- **kurtosis** $E\left\{\left[\frac{X-\mu}{\sigma}\right]^4\right\}$: if $X \sim N(\mu, \sigma^2)$, $Kurt(X) = 3$

II.4.3 Multivariate Distributions: Multinomial Distn

- Consider that a random experiment results in K *mutually exclusive and exhaustive* outcomes $C_1, \ldots, C_K$, with probabilities $\theta_1, \ldots, \theta_K$ for $0 \leq \theta_1, \ldots, \theta_K \leq 1$ and $\sum \theta_k = 1$.
- Repeat the experiment n times independently. Denote the number of kth outcome by X_k for $k = 1, \ldots, K$

Definition. The distribution of rv $(X_1, \ldots, X_K)$ is called a **multinomial** distribution with pmf $p(x_1, \ldots, x_K)$ as follows

$$P(X_1 = x_1, \ldots, X_K = x_K) = \binom{n}{x_1, \ldots, x_K} \theta_1^{x_1} \theta_2^{x_2} \ldots \theta_K^{x_K}$$

for $x_1, \ldots, x_K = 0, 1, \ldots, n$ and $\sum x_k = n$.

II.4.3 Multivariate Distributions: Multinomial Distn

If rv $(X_1, \ldots, X_K)$ follows a multinomial distn $Multin(n; \theta_1, \ldots, \theta_K)$,

- $X_1 + \ldots + X_K = n$ and $\theta_1 + \ldots + \theta_K = 1$
- the special cases of $K = 2$, the binomial distn $B(n, \theta_1)$, and $K = 3$, the trinomial distn $Trino(n; \theta_1, \theta_2, \theta_3)$
- the mgf of $(X_1, \ldots, X_K)$:

$$\left\{\theta_1 e^{t_1} + \ldots + \theta_K e^{t_K}\right\}^n$$

- in general, $X_1 \sim B(n, \theta_1), \ldots, X_K \sim B(n, \theta_K)$
- in general, $(X_1, X_2, Z) \sim Trino(n; \theta_1, \theta_2, \theta_z)$ with $Z = n - X_1 - X_2$, $\theta_z = 1 - \theta_1 - \theta_2$; ... ...

II.4.3 Multivariate Distributions: Multivariate Normal Distn

Definition. Random vector $\mathbf{X} = (X_1, \ldots, X_K)'$ follows a **multivariate normal** distn, denoted by $\mathbf{X} \sim MN(\boldsymbol{\mu}, \Sigma)$, if its pdf is

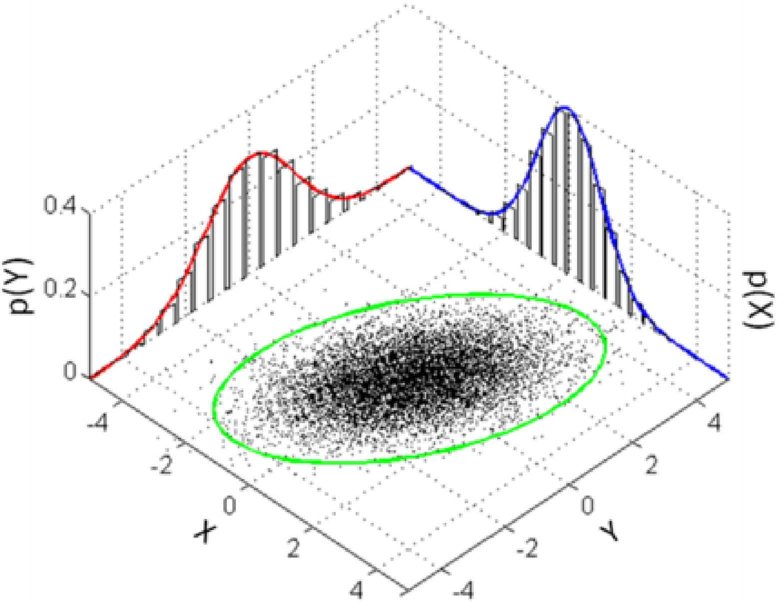
$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{K/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\}, \ \ \textit{for} \ \ \mathbf{x} \in \mathcal{R}^K.$$

Here $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_K)'$ and Σ is a $K \times K$, symmetric, positive semi-definite matrix.

A special case of $K = 2$: rv $(X, Y)'$ follows a **bivariate normal** distn, denoted by $(X, Y) \sim BN(\boldsymbol{\mu}, \Sigma)$, if its pdf (or the joint pdf of X and Y) is, for $x, y \in \mathcal{R}$,

$$f_{X,Y}(x, y) = \frac{1}{(2\pi)^{2/2}|\Sigma|^{1/2}} \exp\left\{ -\frac{1}{2} \begin{pmatrix} x-\mu_1 \\ y-\mu_2 \end{pmatrix}' \Sigma^{-1} \begin{pmatrix} x-\mu_1 \\ y-\mu_2 \end{pmatrix} \right\},$$

II.4.3 Multivariate Distributions: Multivariate Normal Distn



If $\mathbf{X} \sim MN(\boldsymbol{\mu}, \Sigma)$,

- $E(\mathbf{X}) = \boldsymbol{\mu}$, that is $X_k = \mu_k$; $Var(\mathbf{X}) = \Sigma$, that is $Cov(X_i, X_j) = \sigma_{ij}$ if $\Sigma = \left(\sigma_{ij}\right)_{K \times K}$.
- the gmf is $M_{\mathbf{X}}(\mathbf{t}) = E\left[e^{\mathbf{t}'\mathbf{X}}\right] = \exp\left\{\mathbf{t}'\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}'\Sigma\mathbf{t}\right\}$.
- $X_1, \ldots, X_K$ are indpt iff Σ is diagonal.
- rv $Y = a_1X_1 + \ldots + a_KX_K \sim N(\mu_Y, \sigma_Y^2)$ with $\mu_Y = a_1\mu_1 + \ldots + a_K\mu_K = \mathbf{a}'\boldsymbol{\mu}$ and $\sigma_Y^2 = \mathbf{a}'\Sigma\mathbf{a}$.
  - Special cases: $X_1 \sim N(\mu_1, \sigma_{11}), \ldots, X_K \sim N(\mu_K, \sigma_{KK})$
  - If any a linear combination of $W_1, \ldots, W_n$ is normally distributed, $\mathbf{W} \sim MN(\boldsymbol{\mu}_W, \Sigma_W)$.
- rv $\mathbf{Z} = \Sigma^{-1/2}(\mathbf{X} - \boldsymbol{\mu}) \sim MN(\mathbf{0}, \mathbf{I})$.
  - the pdf of $\mathbf{Z}$ is

$$f_{\mathbf{Z}}(\mathbf{z}) = \left(\frac{1}{2\pi}\right)^{K/2} \exp\left\{-\frac{1}{2}\mathbf{z}'\mathbf{z}\right\} = \prod_{k=1}^{K} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}z_k^2\right\},$$

    for $\mathbf{z} = (z_1, \ldots, z_K)' \in \mathcal{R}^K$.
  - $Z_1, \ldots, Z_K$ are *independent and identically distributed* (iid) with the distn $N(0, 1)$.

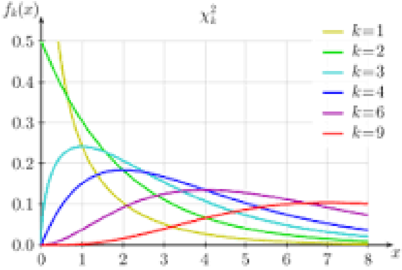
II.4.4 Distributions Induced from Others: χ^2-Distribution

Definition. If $Z_1, \ldots, Z_r$ are indpt and all follow $N(0,1)$, the distribution of $V = Z_1^2 + \ldots + Z_r^2$ is the χ^2**-distribution** with the degrees of freedom (df) r, denoted by $V \sim \chi^2(r)$.

- $E(V) = r$ and $Var(V) = 2r$.
- The pdf of V is

$$f_V(v;r) = \frac{1}{2^{r/2}\Gamma(r/2)} v^{r/2-1} e^{-v/2}, \quad \textit{for } v \geq 0.$$

- $\chi^2(r)$ is a special case of Γ-distn: $\Gamma(r/2, 2)$.

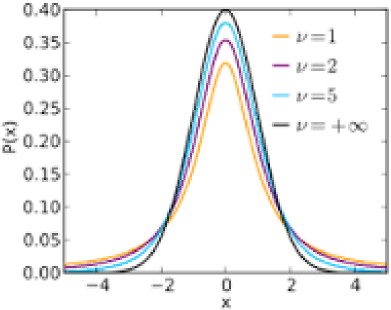


II.4.4 Distributions Induced from Others: Student's t-Distribution

Definition. If $W \sim N(0,1)$, $V \sim \chi^2(r)$, and W, V are indpt, the distn of $T = \frac{W}{\sqrt{V/r}}$ is the **Student's t-distribution** with df r.

(W.S. Gosset, 1908)

- $E(T) = 0$ and $Var(T) = E(T^2) = \frac{r}{r-2}$ if $r > 2$.
- If $r >> 1$, T's distn is approximately $N(0,1)$.
- $Skew(T) = 0$ and $Kurt(T) = \frac{3r-6}{r-4} = 3 + \frac{6}{r-4}$ if $r > 4$.

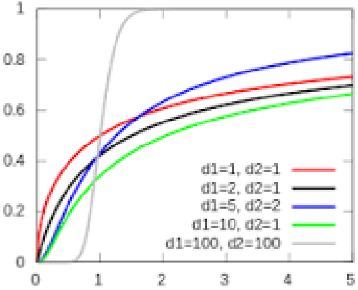


– *Wikipedia*

II.4.4 Distributions Induced from Others: F-Distribution

Definition. If $U \sim \chi^2(r_1)$, $V \sim \chi^2(r_2)$, and U, V are indpt, the distribution of $W = \frac{U/r_1}{V/r_2}$ is the **F-distribution** with the degrees of freedom r_1, r_2, denoted by $W \sim F(r_1, r_2)$.

- $E(W) = \frac{r_2}{r_2-2}$.
- The distn of $1/W$ is $F(r_2, r_1)$.
- If $T \sim t - distn(r)$, $T^2 \sim F(1, r)$.



– *Wikipedia*

What will we do in the next class?

1. Introduction

2. Probability and Distribution (Chp 1-3)

2.4 Some Important Distributions (Chp3)

- *2.4.1 Discrete Distributions*
- *2.4.2 Continuous Distributions*
- *2.4.3 Multivariate Distributions*
- **2.4.4 Distributions Induced from Others**

3. Essential Topics in Mathematical Statistics (Chp 4-6)

- **3.1 Elementary Statistical Inferences**
- *3.2 Consistency and Limiting Distributions*
- *3.3 Maximum Likelihood Methods*

4. Further Topics, Selected from Chp 7-11