MA331 Intermediate Statistics

Lecture 03 Sampling Distributions [1]

Xiaohu Li

Department of Mathematical Sciences
Stevens Institute of Technology
Hoboken, New Jersey 07030

Week 02

[1] Based on Chapter 5.

0. Topics to be covered

This lecture mainly introduces those important statistical distribution concerned with a simple and random sample.

- Sample mean, sample variance and Central Limit Theorem
- Sample proportion, Binomial distribution and Laplace theorem
- χ^2 distribution, Student's t distribution and Snedecor's $\mathcal{F}$ distribtion

1. Population and sample

✌ A population is the collection of all individuals under investigation (denoted by X, Y etc.), and a sample consists of some observations $X_1, \cdots, X_n$ of the population (denoted their realizations by $x_1, \cdots, x_n$).

✌ This course handles the simple and random sample (SRS), in which each observation is randomly selected from the population and not affected by the others. So, X and X_i's are random, and their outcomes are denoted as x and x_i's respectively.

✌ Population distribution: values and probability of the pop's members. Usually it is unknown or partially unknown, thereby we aim to know it based on a sample.

✌ Sample distribution: values and prob for all members of the sample, which are observable.

✌ Statistics: functions of the sample, usually summarizing the sample, their distributions are based on the sampling distribution.

QUERY 1: In R, sample(pop, size, replace, prob) generates a sample of from a finite population.

2. Sampling distributions for count and proportion

✍ **Example**: In a survey of 2500 engineers, 600 of them say they would consider working as a consultant.

- Population: Yes/No answers of all engineers (more than 2500) under study.
- Sample: random variables X_i are observed as $x_i = \begin{cases} 1, & \text{Yes,} \\ 0, & \text{No,} \end{cases}$ and $x_i = 0$ or 1, $i = 1, \cdots, n$, and the sample size $n = 2500$.
- Statistic/Count: total number of 'Y' within the sample (frequency of 'Y')

$$N = \sum_{i=1}^{n} X_i = X_1 + \cdots + X_n, \qquad (0 \le N \le 2500),$$

  where $\sum_{i=1}^{n} x_i = 600$ is observed (a realization of N).
- Statistic/Sample proportion: $\frac{N}{n}$ is the relative frequency of 'Y'.

QUERY 2: Is $X_i = \begin{cases} 1, & \text{Yes,} \\ -1, & \text{No,} \end{cases}$ also a good variable?

3. Binomial distribution — definition

- A number of n observations, all independent of each other.
- Each observation falls into one of two categories: Success or Failure.
- $\mathrm{P}(\text{Success}) = \mathrm{P}(S) = p$ —- information of the population.

✍ The count N of 'S' has Binomial distribution (denoted as $N \sim \mathcal{B}(n, p)$)

$$\mathrm{P}(N = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \cdots, n.$$

✍ Computation of binomial probability $\mathrm{P}(N = 4)$ and $\mathrm{P}(N \leq 4)$: R functions

dbinom(4, 12, 0.2) and pbinom(4, 12, 0.2).

✍ Mean and variance of N:

$$\mu = \mathrm{E}[N] = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = np, \qquad \sigma^2 = \mathrm{Var}[N] = np(1-p).$$

QUERY 3: Have a try to verify the above mean and variance.

4. Binomial distribution — example

✍ 2500 engineers answer Y/N on whether they would like to serve as a consultant.

- Population: Y/N answers of all engineers under study. Specifically, we want to know the population proportion p of 'Y'.
- Sample: random variables X_i observed as $x_i = \begin{cases} 1, & \text{Yes,} \\ 0, & \text{No,} \end{cases}$ and $x_i = 0$ or 1, $i = 1, \cdots, n$, the sample size $n = 2500$.
- Count/frequency of 'Y' within the sample

$$N = \sum_{i=1}^{n} X_i \sim \mathcal{B}(2500, p),$$

here $\sum_{i=1}^{n} x_i = 600$ is observed.

- Sample proportion $\frac{N}{n}$ is a reasonable estimate of p, and here $\frac{N}{n}$ is observed as $600/2500$.

QUERY 4: Say $p \in \{0.1, 0.2, 0.5\}$, based on $\frac{N}{n}$ observed as $\frac{6}{25}$, which p is mostly likely to be true? Why?

5. Laplace theorem – Normal approximation

✍ Due to the difficulty in computing $\binom{n}{k} p^k (1-p)^{n-k}$, Laplace proved the following theorem:

Laplace theorem

For $N \sim \mathcal{B}(n, p)$ with both $np \geq 10$ and $n(1-p) \geq 10$,

$$\mathrm{P}\left(\frac{N/n - p}{\sqrt{p(1-p)/n}} \leq x \right) \approx \Phi(x) \equiv \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x}{2}} \mathrm{d}x, \qquad \text{for all } x.$$

✍ As a result, for all x,

$$\mathrm{P}(N \leq x) = \mathrm{P}\left(\frac{N/n - p}{\sqrt{p(1-p)/n}} \leq \frac{x/n - p}{\sqrt{p(1-p)/n}} \right) \approx \Phi\left(\frac{x - np}{\sqrt{np(1-p)}} \right).$$

That is, approximately $N \sim \mathcal{N}(np, np(1-p))$ or equivalently, $N/n \sim \mathcal{N}(p, p(1-p)/n)$.

✍ We employ $\frac{N}{n}$ to estimate p and hence denote $\hat{p} = \frac{N}{n}$. In practice, we use the following.

Continuity correction

$$\mathrm{P}(N \leq r) = \mathrm{P}(N \leq r + 0.5) \approx \Phi\left(\frac{r + 0.5 - np}{\sqrt{np(1-p)}} \right), \qquad \text{for any integer } r \geq 0.$$

6. Binomial distribution — example continued

✍ Suppose that 26% of all engineers would like to work as consultants. In a survey, 600 of 2500 engineers said 'Yes'.

- $p = 0.26$, $n = 2500$ and N is observed as 600.
- Sample proportion $\hat{p}$ is observed as $600/2500 = 0.24$.
- Mean
$$\mathrm{E}[\hat{p}] = \mathrm{E}\left[\frac{N}{n}\right] = \frac{\mathrm{E}[N]}{n} = \frac{np}{n} = p.$$
- Variance
$$\mathrm{Var}[\hat{p}] = \mathrm{Var}\left[\frac{N}{n}\right] = \frac{\mathrm{Var}[N]}{n^2} = \frac{npq}{n^2} = \frac{pq}{n}.$$
- The probability to observe $\hat{p} \le 0.2$:
$$\mathrm{P}\left(\frac{N}{n} \le 0.2\right) = \mathrm{P}\left(\frac{N/n - p}{\sqrt{p(1-p)/n}} \le \frac{0.24 - 0.26}{\sqrt{0.26(1-0.26)/2500}}\right) \approx \Phi(-0.1368),$$

  which may be found in the normal table or by R function pnorm(-0.1368,0,1)=0.4455944.

QUERY 5: How to evaluate $\mathrm{P}(\frac{N}{n} > 0.3)$ and $\mathrm{P}(0.2 < \frac{N}{n} \le 0.3)$?

7. Sampling distribution for sample mean – Example

Assume 10 of 40 students in this class report their body weight.

- Population: Body weight X of all students in this class (say 40 students get weights $141, 142, \cdots, 180$).
- Sample: observations $X_1, \cdots, X_n$ with sample size $n = 10$.
- Due to randomness, $(X_1, \cdots, X_n)$ may be any of its $\binom{40}{10}$ possible combinations and hence $\bar{X}$ may take the corresponding average values.
- Sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ has its own distribution: each combination of 10 out of 40 students (e.g., $141, \cdots, 150$ with an average 145.5) gets the probability $\binom{40}{10}^{-1}$.
- In practice we want to get the knowledge of $\mu = (141 + \cdots + 180)/40$, which is usually not observable and hence unknown.
- Since $\bar{X}$ approximates the population mean $\mu = \mathrm{E}[X]$, usually it serves as one reasonable estimator of μ.
- To better understand your result based on the sample, we need the distribution of $\bar{X}$.

8. Distribution of sample mean

✍ From a population X we draw a simple random sample (SRS) $X_1, \cdots, X_n$:

- each one is of the distribution of the population X, and
- all X_i's are mutually independent.

✍ To study the mean of X, we usually consider the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Since the sample is random, $\bar{X}$ is also random.

✍ Two important facts:

- The mean of sample mean equals the population mean.

$$\mu_{\bar{X}} = \mathsf{E}[\bar{X}] = \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n}\mathsf{E}[X_i] = \mathsf{E}[X_1] = \mathsf{E}[X] = \mu_X.$$

- The variance of the sample mean equals $\frac{1}{n}$ of the population variance.

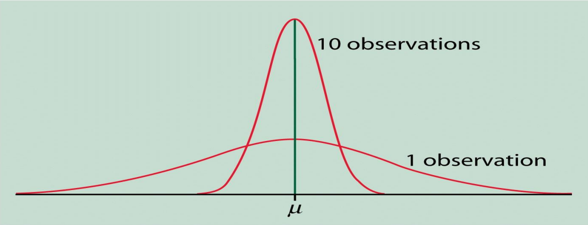
$$\sigma^2_{\bar{X}} = \mathsf{Var}[\bar{X}] = \mathsf{Var}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\mathsf{Var}[X_1] = \frac{\mathsf{Var}[X]}{n} = \frac{\sigma^2_X}{n}.$$

9. Sampling distribution: limit behavior

☞ Since $\bar{X}$ centers on μ_X and the variance goes to 0 as $n \to \infty$, it is reasonable to employ $\bar{X}$ to estimate μ_X and thus we denote $\hat{\mu} = \bar{X}$.

☞ Example: For a sample of size $n = 10$ for the population of the weight X of 42 students, the sample mean $\bar{X}$ has $\binom{42}{n}$ possible outcomes, and $\binom{42}{n} \to 1$ as $n \to 42$.

☞ In general, more and more information about the population is included in the sample as the size grows. Consequently, the randomness/uncertainty in $\bar{X}$ decreases and hence the precision increases.



10. Sampling distribution: Central Limit Theorem (CLT)

☞ $\bar{X} - \mu$, the deviation of $\hat{\mu}$, is random and not accessible. So, we have to study its distribution.

QUERY 6: Why? Do you think it is weird to estimate an unknown number by using a random variable?

✍ For a SRS $X_1, \cdots, X_n$ from $X \sim \mathcal{N}(\mu, \sigma^2)$, since $\bar{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$, it holds that

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1), \qquad \text{for any } n \geq 1.$$

☞ For a SRS $X_1, \cdots, X_n$ from a general (not necessarily normal) population X, the CLT below helps produce an approximation of the precision

☞ Central Limit Theorem

Suppose $-\infty < \mu < +\infty$ and $0 < \sigma^2 < +\infty$. Then, as $n \to \infty$,

$$\mathrm{P}\left(\frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} \leq x\right) = \mathrm{P}\left(\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \leq x\right) \longrightarrow \Phi(x), \quad \text{for all } x.$$

11. Application of CLT in practice

✐ Chase: *Whether dropping the annual fee will increase the amount charged on the credit card?*

✐ The offer is made to 100 customers, the amount charged on cards this year is compared to that next year, and an average increase of \$308 with a standard deviation of \$108 was observed.

✐ Based on the sample $X_1, \cdots, X_n$ of the increase X, the population,

- approximate distribution of the average increase $\bar{X} \sim \mathcal{N}(\mu, \sigma^2/n)$.
- the probability for the average increase to be below \$290:

$$P(\bar{X} \leq 290) = P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq \frac{290-\mu}{\sigma/\sqrt{100}}\right) \approx \Phi\left(\frac{290-\mu}{\sigma/10}\right).$$

- the probability for the average increase to be between \$290 and \$322:

$$P(290 \leq \bar{X} \leq 322) = P(\bar{X} \leq 322) - P(\bar{X} \leq 290) \approx ????.$$

With the knowledge of μ and σ technically we can evaluate those quantities.

12. Some remarks on CLT

✍ Any linear combination of independent normal random variables is also normally distributed. Sometimes the weighted mean

$$w_1 X_1 + \cdots + w_n X_n$$

is also used to estimate the population mean μ.

✐ For a SRS from a population with mean μ and standard deviation σ, when the sample n is large enough, the sampling distribution of $\bar{X}$ is approximately $\mathcal{N}(\mu, \sigma^2/n)$.

✐ What n is large enough? It depends on the population distribution. More observations are required if the population distribution is far away from normal.

- $n = 25$ is generally enough to obtain a normal sampling distribution from a strong skewness or even mild outliers.
- $n = 40$ will typically be good enough to overcome extreme skewness and serious outliers.

13. Univariate normal distribution – definition

✎ A r.v. X is said to be of normal distribution with mean μ and variance σ^2, denoted as $X \sim \mathcal{N}(\mu, \sigma^2)$, if it has the probability density function

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \text{for any } x.$$

✎ $X \sim \mathcal{N}(0, 1)$ is called as standard normal distribution. Specifically, we denote

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, \mathrm{d}x.$$

✎ Normalization

$$X \sim \mathcal{N}(\mu, \sigma^2) \qquad \text{if and only if} \qquad \frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1).$$

Or equivalently,

$$Y \sim \mathcal{N}(0, 1) \qquad \text{if and only if} \qquad \mu + \sigma Y \sim \mathcal{N}(\mu, \sigma^2).$$

✎ The normal random variable $X \sim \mathcal{N}(\mu, \sigma^2)$ has $\mathrm{E}[X] = \mu$ and $\mathrm{Var}[X] = \sigma^2$.

14. Univariate normal distribution – properties

✎ Suppose mutually independent $X_1, \cdots, X_n$ with $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ for $i = 1, \cdots, n$. Then, for real constants $a_1, \cdots, a_n$,

$$a_1 X_1 + \cdots + a_2 X_n = \sum_{i=1}^{n} a_i X_i \sim \mathcal{N}\left(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2 \right).$$

✎ Setting $a_1 = \cdots = a_n = 1/n$, $\mu_1 = \cdots = \mu_n = \mu$ and $\sigma_1^2 = \cdots = \sigma_n^2 = \sigma^2$ in the above, we get

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \sim \mathcal{N}\left(\mu, \frac{1}{n}\sigma^2\right).$$

That is, $\bar{X}$ has the probability density

$$\frac{1}{\sqrt{2\pi\sigma^2/n}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2/n}\right\}, \quad \text{for any } x.$$

✎ In particular, if $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ are independent, then

$$X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2), \qquad X_1 - X_2 \sim \mathcal{N}(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2).$$

15. Multivariate normal distribution

✎ A random vector $\boldsymbol{X} = (X_1, \cdots, X_n)^T$ is said to be of n-dimensional normal distribution with mean vector $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_n)^T$ and covariance metric $\boldsymbol{\Sigma} = (\sigma_{i,j})_{n\times n}$, denoted as $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, if it has the probability density

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\Big\{ -(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\Big\}, \quad \text{for } \boldsymbol{x} = (x_1, \cdots, x_n)^T.$$

✎ For a random vector $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and any real vector $\boldsymbol{a} = (a_1, \cdots, a_n)^T$, the linear combination or inner product is of univariate normal distribution and

$$\boldsymbol{a}^T \boldsymbol{X} = \sum_{i=1}^{n} a_i X_i \sim \mathcal{N}\Big(\boldsymbol{a}^T \boldsymbol{\mu}, \boldsymbol{a}^T \boldsymbol{\Sigma} \boldsymbol{a}\Big).$$

✎ A random vector $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ has

$$(\mathrm{E}[X_1], \cdots, \mathrm{E}[X_n])^T = \mathrm{E}[\boldsymbol{X}] = \boldsymbol{\mu} = (\mu_1, \cdots, \mu_n)^T,$$

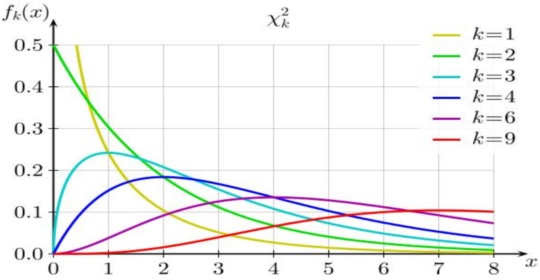
$$\Big(\mathrm{Cov}[X_i, X_j]\Big)_{n\times n} = \mathrm{Cov}[\boldsymbol{X}] = \boldsymbol{\Sigma} = \Big(\sigma_{i,j}\Big)_{n\times n}.$$

16. Three pillars of statistics – χ^2 distribution

✐ Suppose $X_1, \cdots, X_n$ are mutually independent $\mathcal{N}(0,1)$ r.v.'s. Then,

$$X = \sum_{i=1}^{n} X_i^2$$

is said to be of χ_n^2 distribution with degree of freedom (df) n and denoted as $X \sim \chi_n^2$.



✐ The probability density is

$$f_n(x) = \frac{2^{1-\frac{n}{2}} x^{n-1} e^{-\frac{x^2}{2}}}{\Gamma(\frac{n}{2})}, \quad x \geq 0,$$

where gamma function $\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} \mathrm{d}x$, for $t \geq 0$.

✐ Two remarks:

- For a SRS of $\mathcal{N}(\mu, \sigma^2)$, $(n-1)S^2/\sigma^2 \sim \chi_{n-1}^2$.
- In R, distribution function is pchisq(x,n) and quantile function is qchisq(p,n).

- The mean of $X \sim \chi^2_n$:

$$\mathrm{E}[X] = \mathrm{E}\left[\sum_{i=1}^n X_i^2\right] = \sum_{i=1}^n \mathrm{E}[X_i^2] = \sum_{i=1}^n (\mathrm{Var}[X_i] + \mathrm{E}^2[X_i]) = \sum_{i=1}^n (1 + 0^2) = n.$$

- The 4th moment of $X_1 \sim \mathcal{N}(0,1)$:

$$\begin{aligned}
\mathrm{E}[X_1^4] &= \int_{-\infty}^{\infty} x^4 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,\mathrm{d}x = 2\int_0^{\infty} x^4 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,\mathrm{d}x = 2^{3/2} 2\int_0^{\infty} \left(\frac{x^2}{2}\right)^{3/2} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,\mathrm{d}\frac{x^2}{2} \\
&= \frac{2^{5/2}}{\sqrt{2\pi}} \int_0^{\infty} y^{3/2} e^{-y}\,\mathrm{d}y = \frac{4}{\sqrt{\pi}}\left(-y^{3/2} e^{-y}\Big|_0^{\infty} + \frac{3}{2}\int_0^{\infty} y^{1/2} e^{-y}\,\mathrm{d}y\right) \\
&= \frac{2\cdot 3}{\sqrt{\pi}}\left(-y^{1/2} e^{-y}\Big|_0^{\infty} + \frac{1}{2}\int_0^{\infty} y^{-1/2} e^{-y}\,\mathrm{d}y\right) = \frac{3}{\sqrt{\pi}}\Gamma\left(\frac{1}{2}\right) = 3.
\end{aligned}$$

- The variance of $X \sim \chi^2_n$:

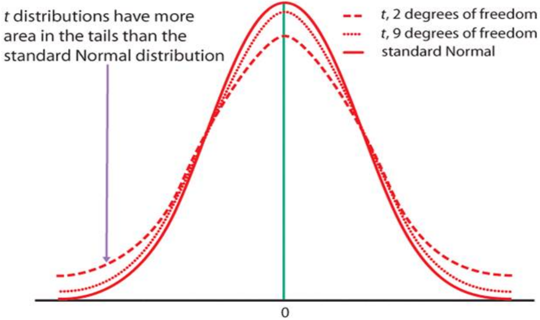
$$\begin{aligned}
\mathrm{Var}[X] = \mathrm{Var}\left[\sum_{i=1}^n X_i^2\right] &= \sum_{i=1}^n \mathrm{Var}[X_i^2] \\
&= \sum_{i=1}^n \left(\mathrm{E}\left[(X_i^2)^2\right] - \mathrm{E}^2[X_i^2]\right) = \sum_{i=1}^n (3 - 1^2) = 2n.
\end{aligned}$$

18. Three pillars of statistics – Student's t distributions

✐ Suppose $X \sim \mathcal{N}(0,1)$ is independent of $Y \sim \chi^2_n$. Then,

$$T = \frac{X}{\sqrt{Y/n}}$$

is said to be of t_n distribution with degree of freedom (df) n and denoted as $T \sim t_n$.



✐ The probability density is

$$f(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}, \ x \geq 0,$$

where $\Gamma(t) = \int_0^\infty x^{t-1}e^{-x}\mathrm{d}x$, for $t \geq 0$.

✐ Three remarks:

- It can be proved that $\mathrm{E}[T] = 0$ and $\mathrm{Var}[T] = \frac{n}{n-2}$.
- For a SRS of $\mathcal{N}(\mu, \sigma^2)$, $\frac{\bar{X}-\mu}{\sqrt{S^2/n}} \sim t_{n-1}$.
- In R, distribution function is pt(x,n) and quantile function is qt(p,n).

19. Three pillars of statistics – F distributions

✐ Suppose $X \sim \chi^2_n$ and $Y \sim \chi^2_m$ are mutually independent r.v.'s. Then,

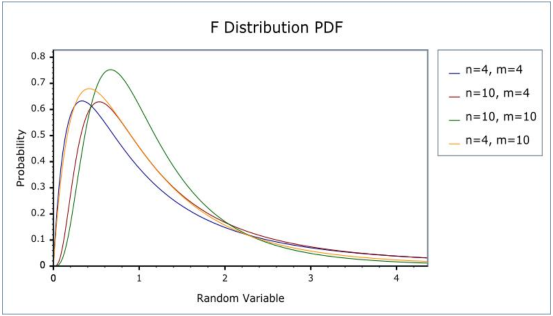
$$F = \frac{X/n}{Y/m}$$

is said to be of $\mathcal{F}_{n,m}$ distribution with degree of freedom (n, m) and denoted as $F \sim \mathcal{F}_{n,m}$.

✐ The probability density is

$$f(x) = \frac{(n/m)^{n/2}}{B(\frac{n}{2}, \frac{m}{2})} x^{\frac{n}{2}-1} (1 + \frac{n}{m} x)^{-\frac{n+m}{2}},$$

where $x \geq 0$ and beta function $B(s,t) = \int_0^1 x^{s-1}(1-x)^{t-1} \mathrm{d}x$, for $s, t \geq 0$.



✐ Three remarks:

- It holds that $\mathrm{E}[X] = \frac{m}{m-2}$ for $m > 2$.
- For independent S_1^2 from $\mathcal{N}(\mu_1, \sigma^2)$ and S_2^2 from $\mathcal{N}(\mu_2, \sigma^2)$, $S_1^2/S_2^2 \sim \mathcal{F}_{n_1-1, n_2-1}$.
- In R, distribution function is pf(x,n,m) and quantile function is qf(p,n,m).

20. Fundamental theorem of sampling distribution

✐ Suppose $(X_1, \cdots, X_n)$ is a SRS of the population $X \sim \mathcal{N}(\mu, \sigma^2)$. Denote the sample mean and sample variance as

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

Then, the following three statements are true.

✌ Fundamental theorem

- The sample mean $\bar{X} \sim \mathcal{N}(\mu, \sigma^2/n)$;
- The sample variance satisfies

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma^2} \sim \chi^2_{n-1};$$

- $\bar{X}$ and S^2 are mutually independent.

➠ The proof involves advanced probability theory and hence is omitted here.