7. Sampling

7.1. Samples, random samples, and statistics

- Let $\tilde{X} = (\tilde{x}_1, ..., \tilde{x}_n)$ be a random vector with distribution $P_{\tilde{X}} : \mathcal{B}(\mathbb{R}^n) \longrightarrow [0, 1]$.
- In statistical inference, a sample is a random vector $\tilde{X} = (\tilde{x}_1, ..., \tilde{x}_n)$.
- To observe a sample of the random vector $\tilde{X}$ is to observe a realization (a value) $X = (x_1, ..., x_n) \in \mathbb{R}^n$ of the random vector $\tilde{X} = (\tilde{x}_1, ..., \tilde{x}_n)$.

- **Definition.** A random sample of size n is a collection of random variables $\{\tilde{x}_i\}_{i=1}^n$ (or $(\tilde{x}_1, ..., \tilde{x}_n)$) that are i.i.d. Its common distribution $P_{\tilde{x}} = P_{\tilde{x}_i}$, for all i, is called the population (or parent) distribution.

- We say that $\{\tilde{x}_i\}_{i=1}^n$ is a random sample of size n from a population $\tilde{x}$ with distribution $P_{\tilde{x}}$.

- Let $\{\tilde{x}_i\}_{i=1}^n$ be a random sample of size n from a population with distribution $P_{\tilde{x}}$, and consider the vector $\tilde{X} = (\tilde{x}_1, ..., \tilde{x}_n)$, then

$$P_{\tilde{X}} = P_{\tilde{x}} \times P_{\tilde{x}} \times ... \times P_{\tilde{x}} \equiv (P_{\tilde{x}})^n$$

- The random vector $\tilde{u} = h\left(\tilde{x}_1, ..., \tilde{x}_n\right)$, where $h : \left(\mathbb{R}^n, \mathcal{B}\left(\mathbb{R}^n\right)\right) \longrightarrow \left(\mathbb{R}^m, \mathcal{B}\left(\mathbb{R}^m\right)\right)$ is a Borel measurable function, is called a **statistic** of the sample $\left\{\tilde{x}_i\right\}_{i=1}^n$.
- The value $u = h\left(x_1, ..., x_n\right) \in \mathbb{R}^m$, where $\left(x_1, ..., x_n\right) \in \mathbb{R}^n$ is a realization of (or a value taken on by) the sample $\left(\tilde{x}_1, ..., \tilde{x}_n\right)$, is the value of the statistic $\tilde{u}$.

- **Two examples of statistics:**

  Let $\{\tilde{x}_i\}_{i=1}^{n}$ be a random sample, then

$$\bar{\mathbf{x}}_n = \bar{\mathbf{x}} = \frac{\sum_{i=1}^{n} \tilde{x}_i}{n}$$

  is the **sample mean** (or mean of the random sample), and

$$\mathbf{s}_n^2 = \mathbf{s}^2 = \frac{\sum_{i=1}^{n} \left(\tilde{x}_i - \bar{\mathbf{x}}_n\right)^2}{n-1}$$

  is the **sample variance** (or variance of the random sample).

- We can omit the subindex n in $\bar{\mathbf{x}}_n$ and $\mathbf{s}_n^2$ when the sample size is fixed.

- Statistics are random variables (or vectors if $m > 1$) and the distribution of a statistic is called "sampling distribution".
- We should distinguish the random variables "statistic", sample mean, or sample variance, $\tilde{u}$, $\bar{\mathbf{x}}_n$, and $\mathbf{s}_n^2$, from the values u, $\overline{x}_n$, and s_n^2 taken by the corresponding random variables.
- If statistics are used to estimate the parameter vector $\theta \in \mathbb{R}^K$ characterizing the distribution $P_{\tilde{X}}(\cdot;\theta)$ of the vector $\tilde{X}$ or the population distribution $P_{\tilde{x}}(\cdot;\theta)$, then they are called "estimators".

7.2. The distribution of the sample mean

- **Theorem.** If $\{\tilde{x}_i\}_{i=1}^n$ is a random sample from a population with the mean μ and the variance σ^2, with $0 < \sigma^2 < \infty$, then

  (a)
  $$\mathrm{E}(\bar{\mathbf{x}}_n) = \mu \quad \text{and} \quad \mathrm{Var}(\bar{\mathbf{x}}_n) = \frac{\sigma^2}{n}.$$

  (b) Strong law of large numbers:
  $$\bar{\mathbf{x}}_n \xrightarrow{a.s} \mu.$$

  (c) Central limit theorem:
  $$\tilde{z}_n \equiv \frac{\bar{\mathbf{x}}_n - \mathrm{E}(\bar{\mathbf{x}}_n)}{\sqrt{\mathrm{Var}(\bar{\mathbf{x}}_n)}} = \frac{\bar{\mathbf{x}}_n - \mu}{\sigma / \sqrt{n}} \overset{a}{\sim} \mathrm{N}(0,1) \qquad (\text{or} \quad \tilde{z}_n \longrightarrow \mathrm{N}(0,1))$$

  or, equivalently,
  $$\sqrt{n}\,(\bar{\mathbf{x}}_n - \mu) \longrightarrow \mathrm{N}(0, \sigma^2).$$

- **Theorem.** If $\bar{\mathbf{x}}_n$ is the mean of a random sample $\{\tilde{x}_i\}_{i=1}^n$ of size n from a normal population $\tilde{x}$ with the mean μ and the variance σ^2, its sampling distribution is a normal distribution with the mean μ and the variance σ^2/n.
- **Proof.**

$$M_{\bar{\mathbf{x}}_n}(t) = M_{\frac{\sum_{i=1}^n \tilde{x}_i}{n}}(t) = M_{\sum_{i=1}^n \tilde{x}_i}\left(\frac{t}{n}\right) = \left[M_{\tilde{x}}\left(\frac{t}{n}\right)\right]^n$$

$$= \left[e^{\mu \frac{t}{n} + \frac{1}{2}\sigma^2 \frac{t^2}{n^2}}\right]^n = e^{\mu t + \frac{1}{2}\frac{\sigma^2}{n}t^2},$$

which is the moment-generating function of a normal distribution with the mean μ and the variance σ^2/n. Therefore,

$\bar{\mathbf{x}}_n \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$. *Q.E.D.*

- Note that

$$\bar{\mathbf{x}}_n \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right) \Longleftrightarrow \frac{\bar{\mathbf{x}}_n - \mu}{\sigma/\sqrt{n}} \sim \mathrm{N}(0,1) \Longleftrightarrow \sqrt{n}(\bar{\mathbf{x}}_n - \mu) \sim \mathrm{N}(0,\sigma^2).$$

7.3. The distribution of the variance of a random sample and the chi-square distribution

- Let $\{\tilde{x}_i\}_{i=1}^n$ be a random sample of size n from a population with the mean μ and the finite variance σ^2.
- $\mathbf{s^2} = \dfrac{\sum_{i=1}^{n} (\tilde{x}_i - \bar{\mathbf{x}})^2}{n-1}$ (sample variance or variance of the random sample).
- $\hat{\mathbf{s}}^{\mathbf{2}} = \dfrac{\sum_{i=1}^{n} (\tilde{x}_i - \bar{\mathbf{x}})^2}{n} = \dfrac{(n-1)\mathbf{s^2}}{n}$.
- $\check{\mathbf{s}}^{\mathbf{2}} = \dfrac{\sum_{i=1}^{n} (\tilde{x}_i - \mu)^2}{n}$.

- **Proposition.** (a) $\mathrm{E}\left(\mathbf{s^2}\right)=\sigma^2$, (b) $\mathrm{E}\left(\hat{\mathbf{s}}^{\mathbf{2}}\right)=\left(\dfrac{n-1}{n}\right)\sigma^2$, and (c) $\mathrm{E}\left(\check{\mathbf{s}}^{\mathbf{2}}\right)=\sigma^2$.

- **Proof. (a)**

$$\mathrm{E}\left(\mathbf{s^2}\right)=\mathrm{E}\left[\frac{\sum_{i=1}^{n}\left(\tilde{x}_i-\bar{\mathbf{x}}\right)^2}{n-1}\right]=\frac{1}{n-1}\mathrm{E}\left(\sum_{i=1}^{n}\left[\left(\tilde{x}_i-\mu\right)-\left(\bar{\mathbf{x}}-\mu\right)\right]^2\right)$$

$$=\frac{1}{n-1}\mathrm{E}\left(\sum_{i=1}^{n}\left[\left(\tilde{x}_i-\mu\right)^2-2\left(\tilde{x}_i-\mu\right)\left(\bar{\mathbf{x}}-\mu\right)+\left(\bar{\mathbf{x}}-\mu\right)^2\right]\right)$$

$$=\frac{1}{n-1}\mathrm{E}\left(\sum_{i=1}^{n}\left(\tilde{x}_i-\mu\right)^2-2n\left(\bar{\mathbf{x}}-\mu\right)^2+n\left(\bar{\mathbf{x}}-\mu\right)^2\right)$$

$$=\frac{1}{n-1}\left(\sum_{i=1}^{n}\mathrm{E}\left[\left(\tilde{x}_i-\mu\right)^2\right]-n\mathrm{E}\left[\left(\bar{\mathbf{x}}-\mu\right)^2\right]\right)$$

$$=\frac{1}{n-1}\left(n\sigma^2-n\frac{\sigma^2}{n}\right)=\frac{1}{n-1}(n-1)\sigma^2=\sigma^2.$$

(b)

$$\mathrm{E}\left(\hat{\mathbf{s}}^{\mathbf{2}}\right)=\mathrm{E}\left[\frac{(n-1)\mathbf{s}^{\mathbf{2}}}{n}.\right]=\left(\frac{n-1}{n}\right)\mathrm{E}\left(\mathbf{s}^{\mathbf{2}}\right)=\left(\frac{n-1}{n}\right)\sigma^{2}.$$

(c)

$$\mathrm{E}\left(\check{\mathbf{s}}^{\mathbf{2}}\right)=\mathrm{E}\left[\frac{\sum_{i=1}^{n}\left(\tilde{x}_i-\mu\right)^2}{n}\right]=\frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\left(\tilde{x}_i-\mu\right)^2=\frac{1}{n}n\sigma^2=\sigma^2. \quad Q.E.D.$$

- *Note*:

$$\mathrm{E}\left(\hat{\mathbf{s}}^{\mathbf{2}}\right)=\left(\frac{n-1}{n}\right)\sigma^2\neq\sigma^2=\mathrm{E}\left(\mathbf{s}^{\mathbf{2}}\right).$$

However, $\mathrm{E}\left(\hat{\mathbf{s}}_n^{\mathbf{2}}\right)$ converges to σ^2 when $n\rightarrow\infty$.

- **Proposition.** If $\{\tilde{z}_i\}_{i=1}^n$ are i.i.d. random variables and $\tilde{z}_i \sim \mathrm{N}(0,1)$, for $i = 1, 2, ..., n$, then

$$\tilde{y} = \sum_{i=1}^{n} \tilde{z}_i^2 \sim \chi_n^2.$$

- **Proof.** Remember that $\tilde{z}_i^2 \sim \chi_1^2$ so that

$$M_{\tilde{z}_i^2}(t) = (1-2t)^{-1/2}, \quad \text{for } t < 1/2.$$

Therefore, from independency,

$$M_{\tilde{y}}(t) = \left[(1-2t)^{-1/2}\right]^n = (1-2t)^{-n/2}, \quad \text{for } t < 1/2,$$

which is the moment-generating function of a random variable whose distribution is χ_n^2. Thus, $\tilde{y} \sim \chi_n^2$. *Q.E.D.*

- **Proposition.** If $\{\tilde{x}_i\}_{i=1}^n$ are independent random variables and $\tilde{x}_i \sim \chi^2_{\nu_i}$, for $i = 1, 2, ..., n$, then

$$\tilde{y} = \sum_{i=1}^n \tilde{x}_i \sim \chi^2_{(\nu_1+\nu_2+...+\nu_n)}.$$

- **Proof.** From independency,

$$M_{\tilde{y}}(t) = \prod_{i=1}^n (1-2t)^{-\nu_i/2} = (1-2t)^{-(\nu_1+\nu_2+...+\nu_n)/2} \quad \text{for } t < 1/2.$$

Therefore, $\tilde{y} \sim \chi^2_{(\nu_1+\nu_2+...+\nu_n)}$. *Q.E.D.*

- **Proposition.** If $\tilde{x}_1$ and $\tilde{x}_2$ are independent random variables and $\tilde{x}_1 \sim \chi^2_{\nu_1}$ and $\tilde{x}_1 + \tilde{x}_2 \sim \chi^2_{\nu}$, with $\nu > \nu_1$, then $\tilde{x}_2 \sim \chi^2_{(\nu-\nu_1)}$.
- **Proof.** From independency,

$$M_{\tilde{x}_1+\tilde{x}_2}(t) = M_{\tilde{x}_1}(t) \cdot M_{\tilde{x}_2}(t)$$

or

$$(1-2t)^{-\nu/2} = (1-2t)^{-\nu_1/2} M_{\tilde{x}_2}(t), \quad \text{for } t < 1/2.$$

Therefore,

$$M_{\tilde{x}_2}(t) = (1-2t)^{-(\nu-\nu_1)/2}, \quad \text{for } t < 1/2.$$

Thus, $\tilde{x}_2 \sim \chi^2_{(\nu-\nu_1)}$. *Q.E.D.*

- **Theorem.** If $\bar{\mathbf{x}}$ and $\mathbf{s^2}$ are the mean and the variance of a random sample $\{\tilde{x}_i\}_{i=1}^n$ of size n from a normal population with mean μ and variance σ^2, then

  (1) $\bar{\mathbf{x}}$ and $\mathbf{s^2}$ are independent.

  (2)
$$\frac{(n-1)\mathbf{s^2}}{\sigma^2} \equiv \frac{\sum_{i=1}^{n} \left(\tilde{x}_i - \bar{\mathbf{x}}\right)^2}{\sigma^2} \sim \chi^2_{n-1}.$$

- **Proof. (1)**

  (a) $\tilde{x}_n - \bar{\mathbf{x}} = -\sum_{i=1}^{n-1} \left(\tilde{x}_i - \bar{\mathbf{x}}\right)$ since $\sum_{i=1}^{n} \tilde{x}_i = n\bar{\mathbf{x}}$ (or $\sum_{i=1}^{n} \left(\tilde{x}_i - \bar{\mathbf{x}}\right) = 0$);

  (b) $\mathbf{s^2} = \dfrac{\sum_{i=1}^{n} \left(\tilde{x}_i - \bar{\mathbf{x}}\right)^2}{n-1}$ is thus a function of the random vector $\widetilde{w} = \left(\tilde{x}_1 - \bar{\mathbf{x}}, \tilde{x}_2 - \bar{\mathbf{x}}, ..., \tilde{x}_{n-1} - \bar{\mathbf{x}}\right)$;

  (c) The n-dimensional random vector $\left(\bar{\mathbf{x}}, \widetilde{w}\right)^{\mathsf{T}}$ has a multivariate normal distribution according to the General Proposition in Section 4.10 of Chapter 4 since

$$(\bar{\mathbf{x}}, \widetilde{w})^{\mathsf{T}} = \begin{pmatrix} \bar{\mathbf{x}} \\ \tilde{x}_1 - \bar{\mathbf{x}} \\ \tilde{x}_2 - \bar{\mathbf{x}} \\ \vdots \\ \tilde{x}_{n-1} - \bar{\mathbf{x}} \end{pmatrix}_{n\times 1} =$$

$$\begin{pmatrix} \frac{1}{n} & \frac{1}{n} & \frac{1}{n} & \vdots & \frac{1}{n} & \frac{1}{n} \\ 1-\frac{1}{n} & -\frac{1}{n} & -\frac{1}{n} & \vdots & -\frac{1}{n} & -\frac{1}{n} \\ -\frac{1}{n} & 1-\frac{1}{n} & -\frac{1}{n} & \vdots & -\frac{1}{n} & -\frac{1}{n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -\frac{1}{n} & -\frac{1}{n} & -\frac{1}{n} & \vdots & 1-\frac{1}{n} & -\frac{1}{n} \end{pmatrix}_{n\times n} \times \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \vdots \\ \tilde{x}_n \end{pmatrix}_{n\times 1}$$

and $(\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n)^\mathsf{T}$ is multivariate normal because it is a vector of independent, normally distributed variables;

(d) $\mathrm{Cov}(\bar{\mathbf{x}}, \tilde{x}_i - \bar{\mathbf{x}}) = 0$ for all i (Exercise) so that $\bar{\mathbf{x}}$ and $\widetilde{w}$ are uncorrelated and, thus, they are independent. Therefore, $\bar{\mathbf{x}}$ and $\mathbf{s^2}$ are independent.

- **(2)** The following equality holds:

$$\sum_{i=1}^{n} \left(\tilde{x}_i - \mu\right)^2 = \sum_{i=1}^{n} \left(\tilde{x}_i - \bar{\mathbf{x}}\right)^2 + n\left(\bar{\mathbf{x}} - \mu\right)^2 \qquad (*)$$

since it is equivalent to

$$\sum_{i=1}^{n} \left(\tilde{x}_i^2 - 2\tilde{x}_i\mu + \mu^2\right) = \sum_{i=1}^{n} \left(\tilde{x}_i^2 - 2\tilde{x}_i\bar{\mathbf{x}} + \bar{\mathbf{x}}^2\right) + n\left(\bar{\mathbf{x}}^2 - 2\bar{\mathbf{x}}\mu + \mu^2\right).$$

As $\sum_{i=1}^{n} \tilde{x}_i = n\bar{\mathbf{x}}$, the previous expression becomes

$$\sum_{i=1}^{n} \tilde{x}_i^2 - 2n\bar{\mathbf{x}}\mu + n\mu^2 = \sum_{i=1}^{n} \tilde{x}_i^2 \underbrace{-2n\bar{\mathbf{x}}^2 + n\bar{\mathbf{x}}^2 + n\bar{\mathbf{x}}^2}_{=0} - 2n\bar{\mathbf{x}}\mu + n\mu^2.$$

Divide $(*)$ by σ^2,

$$\sum_{i=1}^{n} \left(\frac{\tilde{x}_i - \mu}{\sigma}\right)^2 = \frac{\sum_{i=1}^{n} \left(\tilde{x}_i - \bar{\mathbf{x}}_n\right)^2}{\sigma^2} + \left(\frac{\bar{\mathbf{x}} - \mu}{\sigma / \sqrt{n}}\right)^2.$$

- We know that $\dfrac{\tilde{x}_i - \mu}{\sigma} \sim \mathrm{N}(0,1)$ so that $\displaystyle\sum_{i=1}^{n} \left(\frac{\tilde{x}_i - \mu}{\sigma} \right)^2 \sim \chi_n^2$ and $\dfrac{\bar{\mathbf{x}} - \mu}{\sigma / \sqrt{n}} \sim \mathrm{N}(0,1)$ so that $\left(\dfrac{\bar{\mathbf{x}} - \mu}{\sigma / \sqrt{n}} \right)^2 \sim \chi_1^2$. Therefore,

  $$\frac{\sum_{i=1}^{n} (\tilde{x}_i - \bar{\mathbf{x}})^2}{\sigma^2} \equiv \frac{(n-1)\mathbf{s^2}}{\sigma^2} \sim \chi_{n-1}^2 \text{ since } \left(\frac{\bar{\mathbf{x}} - \mu}{\sigma / \sqrt{n}} \right)^2 \text{ and } \frac{(n-1)\mathbf{s^2}}{\sigma^2}$$

  are independent as follows from part (a). *Q.E.D.*

- Note that only $n-1$ terms of the sum $\displaystyle\sum_{i=1}^{n} (\tilde{x}_i - \bar{\mathbf{x}})^2$ are allowed to vary freely since they are constrained by the relation $\displaystyle\sum_{i=1}^{n} (\tilde{x}_i - \bar{\mathbf{x}}) = 0$.

  This is why we say that the statistic $\dfrac{(n-1)\mathbf{s^2}}{\sigma^2}$ has $n-1$ degrees of freedom. In general, the number of degrees of freedom is the number of values in the final calculation of a statistic that are free to vary.

- **Proposition.** Let $\{\tilde{x}_i\}_{i=1}^n$ be a random sample of size n from a normal population with mean μ and variance σ^2. Then,

  (a) $\mathrm{Var}\left(\mathbf{s^2}\right) = \dfrac{2\sigma^4}{n-1}$, (b) $\mathrm{Var}\left(\hat{\mathbf{s}}^{\mathbf{2}}\right) = \dfrac{2(n-1)\sigma^4}{n^2}$, and

  (c) $\mathrm{Var}\left(\check{\mathbf{s}}^{\mathbf{2}}\right) = \dfrac{2\sigma^4}{n}$.

- **Proof. (a)** Since

$$\frac{\sum_{i=1}^{n} \left(\tilde{x}_i - \bar{\mathbf{x}}\right)^2}{\sigma^2} = \frac{(n-1)\mathbf{s^2}}{\sigma^2} \sim \chi^2_{n-1},$$

  we get

$$\mathrm{Var}\left[\frac{(n-1)\mathbf{s^2}}{\sigma^2}\right] = \frac{(n-1)^2}{\sigma^4}\mathrm{Var}\left(\mathbf{s^2}\right) = 2\left(n-1\right)$$

  and, thus, $\mathrm{Var}\left(\mathbf{s^2}\right) = \dfrac{2\sigma^4}{n-1}$.

(b)

$$\operatorname{Var}\left(\hat{\mathbf{s}}^{\mathbf{2}}\right)=\operatorname{Var}\left[\frac{(n-1)\mathbf{s}^{\mathbf{2}}}{n}\right]=\left(\frac{n-1}{n}\right)^{2}\operatorname{Var}\left(\mathbf{s}^{\mathbf{2}}\right)$$

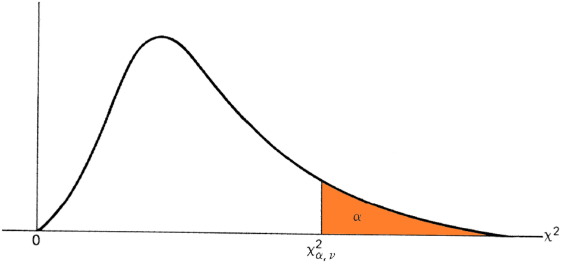
$$=\left(\frac{n-1}{n}\right)^{2}\frac{2\sigma^{4}}{n-1}=\frac{2(n-1)\sigma^{4}}{n^{2}}.$$

(c) Since $\sum_{i=1}^{n}\left(\frac{\tilde{x}_i-\mu}{\sigma}\right)^{2}\sim\chi_n^2$, we get

$$\operatorname{Var}\left[\sum_{i=1}^{n}\left(\frac{\tilde{x}_i-\mu}{\sigma}\right)^{2}\right]=\frac{n^2}{\sigma^4}\operatorname{Var}\left[\frac{\sum_{i=1}^{n}\left(\tilde{x}_i-\mu\right)^{2}}{n}\right]=\frac{n^2}{\sigma^4}\operatorname{Var}\left(\breve{\mathbf{s}}^{\mathbf{2}}\right)=2n$$

and, thus, $\operatorname{Var}\left(\breve{\mathbf{s}}^{\mathbf{2}}\right)=\frac{2\sigma^4}{n}$. *Q.E.D.*

- In the table of the chi-square we find the value $\chi^2_{\alpha,\nu}$ such that
$$P\left\{\tilde{x} \geq \chi^2_{\alpha,\nu}\right\} = P_{\tilde{x}}\left[\chi^2_{\alpha,\nu}, \infty\right) = \alpha,$$
when $\tilde{x} \sim \chi^2_\nu$.



- If $\tilde{x}_\nu \sim \chi^2_\nu$, then $\frac{\tilde{x}_\nu - \nu}{\sqrt{2\nu}} \longrightarrow \mathrm{N}(0,1)$ as $\nu \longrightarrow \infty$. (Exercise).
- Thus, $P\left\{\tilde{x}_\nu \geq \chi^2_{\alpha,\nu}\right\} = P\left\{\frac{\tilde{x}_\nu - \nu}{\sqrt{2\nu}} \geq \frac{\chi^2_{\alpha,\nu} - \nu}{\sqrt{2\nu}}\right\} = \alpha \approx 1 - \mathrm{N}\left(\frac{\chi^2_{\alpha,\nu} - \nu}{\sqrt{2\nu}}\right)$ for ν large.
- Hence, if $\tilde{z} \sim \mathrm{N}(0,1)$ and $P\left\{\tilde{z} \geq z_\alpha\right\} = \alpha$, then $\frac{\chi^2_{\alpha,\nu} - \nu}{\sqrt{2\nu}} \approx z_\alpha$ and, thus, $\chi^2_{\alpha,\nu} \approx \nu + \left(\sqrt{2\nu}\right) z_\alpha$ for ν large.

7.4. The t distribution

William S. Gosset ("Student") (1876 – 1937)

- **Definition.** A random variable $\tilde{x}$ has the (Student's) t distribution (or is t) with $\nu > 0$ degrees of freedom if its density is

$$f_{\tilde{x}}(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\cdot\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} \quad \text{for } -\infty < x < \infty.$$

- We write $\tilde{x} \sim \mathrm{t}_\nu$.
- Note that the t density is symmetric with respect to zero.
- **Theorem.** If $\tilde{y}$ and $\tilde{z}$ are independent random variables with $\tilde{y} \sim \chi^2_\nu$ and $\tilde{z} \sim \mathrm{N}(0,1)$, then

$$\tilde{x} = \frac{\tilde{z}}{\sqrt{\tilde{y}\,/\nu}} \sim \mathrm{t}_\nu.$$

- **Proof.** See the handout.

- Let $\tilde{x} \sim \mathrm{t}_\nu$. Then $\tilde{x} \in \mathrm{L}^k$ if and only if $\nu > k$. Thus, the t distribution has not well-defined (i.e., finite) moment-generating function in a neighborhood of 0. In fact, $M_{\tilde{x}}(t)$ is finite only at $t = 0$.

- Moments: If $\nu > k$, then

$$\mu'_k = \begin{cases} \dfrac{\nu^{k/2}\Gamma\left(\frac{k+1}{2}\right)\Gamma\left(\frac{\nu-k}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{\nu}{2}\right)} & \text{if } k \text{ is even} \\ 0 & \text{if } k \text{ is odd.} \end{cases}$$

Mean:

$$\text{If } \nu > 1, \text{ then } \mathrm{E}(\tilde{x}) = 0.$$

Variance:

$$\text{If } \nu > 2, \text{ then } \mathrm{Var}(\tilde{x}) = \frac{\nu}{\nu - 2}.$$

- When $\nu = 1$, the t distribution is the Cauchy distribution with parameters $\alpha = 0$ and $\beta = 1$. Recall that the density of the Cauchy distribution is

$$f_{\tilde{x}}(x) = \frac{\beta/\pi}{(x-\alpha)^2 + \beta^2} \quad \text{for } -\infty < x < \infty,$$

  and that the Cauchy distribution has not well-defined mean (and thus has indefinite/infinite higher odd/even central moments).

- Therefore, the density of a t_1 random variable is

$$f_{\tilde{x}}(x) = \frac{1}{\pi(1+x^2)} \quad \text{for } -\infty < x < \infty. \qquad \longleftarrow \text{Exercise}$$

- If $\tilde{x}_\nu \sim t_\nu$, then $\tilde{x}_\nu \longrightarrow \mathrm{N}(0,1)$ as $\nu \longrightarrow \infty$. (Exercise).

- **Theorem.** If $\bar{\mathbf{x}}$ and $\mathbf{s^2}$ are the mean and the variance of a random sample $\{\tilde{x}_i\}_{i=1}^n$ of size n from a normal population with mean μ and variance σ^2, then

$$\frac{\bar{\mathbf{x}}-\mu}{\mathbf{s}\big/\sqrt{n}} \sim \mathrm{t}_{n-1},$$

  where $\mathbf{s}$ is the sample standard deviation (or standard deviation of the random sample), $\mathbf{s} = \left(\mathbf{s^2}\right)^{1/2}$.

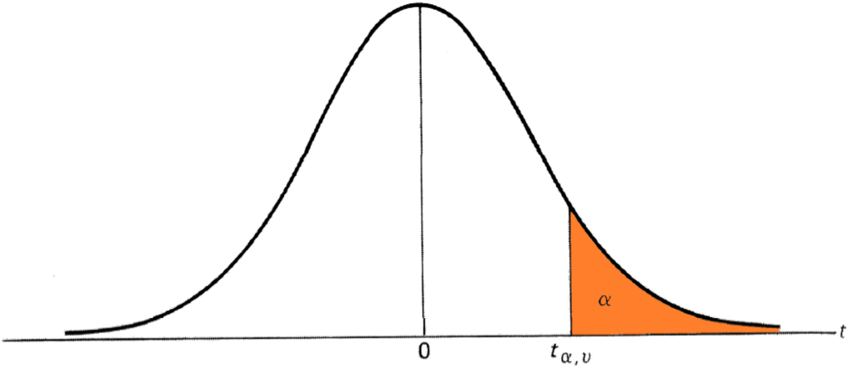
- **Proof.** Let $\tilde{y} = \dfrac{(n-1)\mathbf{s^2}}{\sigma^2} \sim \chi^2_{n-1}$ and $\tilde{z} = \dfrac{\bar{\mathbf{x}}-\mu}{\sigma\big/\sqrt{n}} \sim \mathrm{N}(0,1)$. The random variables $\tilde{z}$ and $\tilde{y}$ are independent. Then,

$$\frac{\dfrac{\bar{\mathbf{x}}-\mu}{\sigma\big/\sqrt{n}}}{\sqrt{\dfrac{(n-1)\mathbf{s^2}}{\sigma^2}\Big/(n-1)}} = \frac{\bar{\mathbf{x}}-\mu}{\mathbf{s}\big/\sqrt{n}} \sim \mathrm{t}_{n-1}. \qquad \textit{Q.E.D.}$$

- In the table of the t_ν we find the value $t_{\alpha,\nu}$ such that

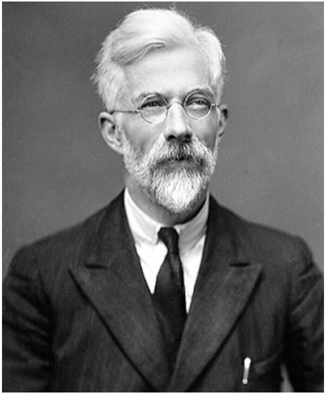
$$P\left\{\tilde{x} \geq t_{\alpha,\nu}\right\} = P_{\tilde{x}}\left[t_{\alpha,\nu}, \infty\right) = \alpha,$$

when $\tilde{x} \sim t_\nu$.



- Note that $P\left\{\tilde{x} \leq -t_{\alpha,\nu}\right\} = \alpha$ or $P\left\{\tilde{x} \geq -t_{\alpha,\nu}\right\} = 1-\alpha$. Moreover, $P\left\{\tilde{x} \leq t_{\alpha,\nu}\right\} = 1-\alpha$.

7.5. The F distribution



Ronald Fisher (1890 - 1962)

George W. Snedecor (1881 - 1974)

- **Definition.** A random variable $\tilde{y}$ has the (Fisher-Snedecor's) F distribution (or is F) with $\nu_1 > 0$ and $\nu_2 > 0$ degrees of freedom if its density is

$$f_{\tilde{y}}(y) = \begin{cases} \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\cdot\Gamma\left(\frac{\nu_2}{2}\right)} \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} y^{\frac{\nu_1}{2}-1} \left(1+\frac{\nu_1}{\nu_2}y\right)^{-\frac{1}{2}(\nu_1+\nu_2)} & \text{for } y > 0 \\ 0 & \text{elsewhere.} \end{cases}$$

- We write $\tilde{y} \sim \mathrm{F}_{\nu_1,\nu_2}$.
- Note that $\mathrm{F}_{\nu_1,\nu_2} \neq \mathrm{F}_{\nu_2,\nu_1}$ but $\tilde{x} = \dfrac{1}{\tilde{y}} \sim \mathrm{F}_{\nu_2,\nu_1}$ (Exercise).
- **Theorem.** If $\tilde{u}$ and $\tilde{v}$ are independent random variables with $\tilde{u} \sim \chi^2_{\nu_1}$ and $\tilde{v} \sim \chi^2_{\nu_2}$, then

$$\tilde{y} = \frac{\tilde{u}/\nu_1}{\tilde{v}/\nu_2} \sim \mathrm{F}_{\nu_1,\nu_2}.$$

- **Proof.** See the handout.

- The F distribution has well-defined (i.e., finite) moment-generating function $M_{\tilde{y}}(t)$ if and only if $t \leq 0$.
- Let $\tilde{y} \sim \mathrm{F}_{\nu_1,\nu_2}$. Then, $\tilde{y} \in \mathrm{L}^k$ if and only if $\nu_2 > 2k$.
- Moments:

$$\mu_k' = \left(\frac{\nu_2}{\nu_1}\right)^k \frac{\Gamma\left(\frac{\nu_1}{2}+k\right)\Gamma\left(\frac{\nu_2}{2}-k\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}, \quad \text{for } \nu_2 > 2k.$$

Mean:

$$\text{If } \nu_2 > 2, \text{ then } \mathrm{E}(\tilde{y}) = \frac{\nu_2}{\nu_2 - 2}.$$

Variance:

$$\text{If } \nu_2 > 4, \text{ then } \mathrm{Var}(\tilde{y}) = \frac{2\nu_2^2\left(\nu_1+\nu_2-2\right)}{\nu_1(\nu_2-2)^2\left(\nu_2-4\right)}.$$

- **Theorem.** Let $\mathbf{s}_1^2$ and $\mathbf{s}_2^2$ be the variances of two independent random samples of size n_1 and n_2 (i.e., the $n_1 + n_2$ random variables constituting the two random samples are independent) from two normal populations with variances σ_1^2 and σ_2^2, respectively. Then,

$$\tilde{y} = \frac{\mathbf{s}_1^2 / \sigma_1^2}{\mathbf{s}_2^2 / \sigma_2^2} = \frac{\sigma_2^2 \cdot \mathbf{s}_1^2}{\sigma_1^2 \cdot \mathbf{s}_2^2} \sim \mathrm{F}_{n_1-1, n_2-1}.$$

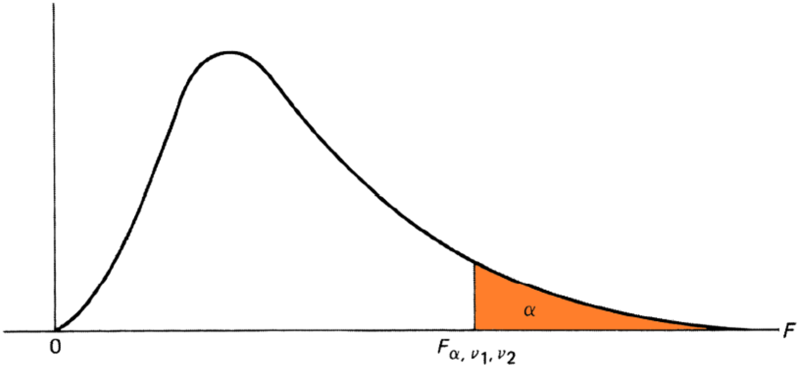
- **Proof.** Let $\tilde{u} = \dfrac{(n_1-1)\mathbf{s}_1^2}{\sigma_1^2} \sim \chi^2_{n_1-1}$ and $\tilde{v} = \dfrac{(n_2-1)\mathbf{s}_2^2}{\sigma_2^2} \sim \chi^2_{n_2-1}$. The random variables $\tilde{u}$ and $\tilde{v}$ are independent. Then

$$\frac{\dfrac{(n_1-1)\mathbf{s}_1^2}{\sigma_1^2} \Big/ (n_1-1)}{\dfrac{(n_2-1)\mathbf{s}_2^2}{\sigma_2^2} \Big/ (n_2-1)} = \frac{\mathbf{s}_1^2 / \sigma_1^2}{\mathbf{s}_2^2 / \sigma_2^2} \sim \mathrm{F}_{n_1-1, n_2-1}. \qquad \textit{Q.E.D.}$$

- In the table of the F_{ν_1,ν_2} we find the value F_{α,ν_1,ν_2} such that

$$P\left\{\tilde{y} \geq F_{\alpha,\nu_1,\nu_2}\right\} = P_{\tilde{y}}\left[F_{\alpha,\nu_1,\nu_2}, \infty\right) = \alpha,$$

when $\tilde{y} \sim \mathrm{F}_{\nu_1,\nu_2}$.



- Note that $F_{\alpha,\nu_1,\nu_2} = \dfrac{1}{F_{1-\alpha,\nu_2,\nu_1}}$ (think about it).

7.6. Order statistics

- The jth order statistic $\tilde{x}_{(j)}$ of a random sample $\{\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n\}$ of size n is the random variable of the sample that takes the jth smallest value.
- See the handout for the distribution of the order statistics.

The t Distribution

Theorem. If $\widetilde{y}$ and $\widetilde{z}$ are independent random variables with $\widetilde{y} \sim \chi^2_\nu$ and $\widetilde{z} \sim \mathrm{N}(0,1)$, then

$$\widetilde{x} = \frac{\widetilde{z}}{\sqrt{\widetilde{y}\,/\nu}} \sim \mathrm{t}_\nu .$$

Proof. Since $\tilde{z}$ and $\tilde{y}$ are independent, their joint density is given by

$$f_{\widetilde{y},\widetilde{z}}(y,z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \frac{1}{\Gamma\left(\dfrac{\nu}{2}\right) 2^{\nu/2}} y^{\frac{\nu}{2}-1} e^{-y/2}$$

for $y > 0$, $-\infty < z < \infty$ and $f_{\widetilde{y},\widetilde{z}}(y,z) = 0$ elsewhere. Then, we solve $x = \dfrac{z}{\sqrt{y/\nu}}$ for z, getting $z = x\sqrt{y/\nu}$. Thus, we use the following change of variable:

$$g^{-1} : \begin{cases} z = x\sqrt{y/\nu} \\ \\ y = y \end{cases} \Rightarrow |J_{g^{-1}}| = \sqrt{y/\nu} \quad ; \qquad g : \begin{cases} x = \dfrac{z}{\sqrt{y/\nu}} \in (-\infty, \infty) \\ \\ y = y > 0 \end{cases}$$

Therefore, the joint density of $\tilde{y}$ and $\tilde{x}$ is given by (check it!):

$$f_{\widetilde{y},\widetilde{x}}(y,x) = \begin{cases} \dfrac{1}{\sqrt{2\pi\nu} \cdot \Gamma\left(\dfrac{\nu}{2}\right) 2^{\nu/2}} y^{\frac{\nu-1}{2}} e^{-\frac{y}{2}\left(1+\frac{x^2}{\nu}\right)} & \text{for } y > 0,\ x \in (-\infty,\infty) \\ \\ 0 & \text{elsewhere} \end{cases}$$

and integrating out y with the aid of the change of variable $w = \frac{y}{2}\left(1 + \frac{x^2}{\nu}\right)$, which implies that $\frac{dy}{dw} = \dfrac{2}{1+\frac{x^2}{\nu}}$, we get the following density (check it!):

$$f_{\tilde{x}}(x) = \frac{\Gamma\left(\dfrac{\nu+1}{2}\right)}{\sqrt{\pi\nu} \cdot \Gamma\left(\dfrac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad \text{for } x \in (-\infty, \infty),$$

which is the t density with ν degrees of freedom. *Q.E.D.*

– – – – – – – – – – – – – – –

The F Distribution

Theorem. If $\widetilde{u}$ and $\widetilde{v}$ are independent random variables with $\widetilde{u} \sim \chi^2_{\nu_1}$ and $\widetilde{v} \sim \chi^2_{\nu_1}$, then

$$\widetilde{y} = \frac{\widetilde{u}/\,\nu_1}{\widetilde{v}/\,\nu_2} \sim \mathrm{F}_{\nu_1,\nu_2}.$$

Proof. Since $\tilde{u}$ and $\tilde{v}$ are independent, the joint density of $\tilde{u}$ and $\tilde{v}$ is given by

$$\begin{aligned} f_{\widetilde{u},\widetilde{v}}(u,v) &= \frac{1}{2^{\nu_1/2}\cdot\Gamma(\frac{\nu_1}{2})} u^{\frac{\nu_1}{2}-1} e^{-\frac{u}{2}} \frac{1}{2^{\nu_2/2}\Gamma(\frac{\nu_2}{2})} v^{\frac{\nu_2}{2}-1} e^{-v/2} \\ &= \frac{1}{2^{(\nu_1+\nu_2)/2}\cdot\Gamma(\frac{\nu_1}{2})\cdot\Gamma(\frac{\nu_2}{2})} u^{\frac{\nu_1}{2}-1} v^{\frac{\nu_2}{2}-1} e^{-\frac{(u+v)}{2}}, \end{aligned}$$

for $u > 0$, $v > 0$ and $f_{\widetilde{u},\widetilde{v}}(u,v) = 0$, elsewhere. Then, we solve $y = \dfrac{u/\nu_1}{v/\nu_2}$ for u, getting $u = \dfrac{\nu_1}{\nu_2} vy$. Thus, we use the following change of variable

$$g^{-1}: \begin{cases} u = \dfrac{\nu_1}{\nu_2} vy \\ \\ v = v \end{cases} \Rightarrow |J_{g^{-1}}| = \frac{\nu_1}{\nu_2} v \quad ; \qquad g: \begin{cases} y = \dfrac{u\,/\nu_1}{v\,/\nu_2} > 0 \\ \\ v = v > 0 \end{cases}$$

Therefore, the joint density of $\tilde{y}$ and $\tilde{v}$ is given by

$$f_{\widetilde{y},\widetilde{v}}(y,v) = \frac{\left(\dfrac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}}}{2^{(\nu_1+\nu_2)/2}\cdot\Gamma\left(\dfrac{\nu_1}{2}\right)\cdot\Gamma\left(\dfrac{\nu_2}{\nu_1}\right)} y^{\frac{\nu_1}{2}-1} v^{\frac{\nu_1+\nu_2}{2}-1} e^{-\frac{v}{2}(\frac{\nu_1 y}{\nu_2}+1)}$$

for $y > 0$ and $v > 0$; and $f_{\widetilde{y},\widetilde{v}}(y,v) = 0$ elsewhere (check it!). Now integrating out v by using the change of variable $w = \dfrac{v}{2}\left(\dfrac{\nu_1 y}{\nu_2} + 1\right)$, we finally get the following density (check it!):

$$f_{\widetilde{y}}(y) = \begin{cases} \dfrac{\Gamma\left(\dfrac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\dfrac{\nu_1}{2}\right)\cdot\Gamma\left(\dfrac{\nu_2}{2}\right)} \left(\dfrac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} y^{\frac{\nu_1}{2}-1} \left(1 + \dfrac{\nu_1}{\nu_2} y\right)^{-\frac{1}{2}(\nu_1+\nu_2)} & \text{for } y > 0 \\ \\ 0 & \text{elsewhere,} \end{cases}$$

which is the F density with ν_1 and ν_2 degrees of freedom. *Q.E.D.*

The Chi-square distribution

Values of $\chi^2_{\alpha,\nu}$

ν	α = .995	α = .99	α = .975	α = .95	α = .05	α = .025	α = .01	α = .005	ν
1	.0000393	.000157	.000982	.00393	3.841	5.024	6.635	7.879	1
2	.0100	.0201	.0506	.103	5.991	7.378	9.210	10.597	2
3	.0717	.115	.216	.352	7.815	9.348	11.345	12.838	3
4	.207	.297	.484	.711	9.488	11.143	13.277	14.860	4
5	.412	.554	.831	1.145	11.070	12.832	15.086	16.750	5
6	.676	.872	1.237	1.635	12.592	14.449	16.812	18.548	6
7	.989	1.239	1.690	2.167	14.067	16.013	18.475	20.278	7
8	1.344	1.646	2.180	2.733	15.507	17.535	20.090	21.955	8
9	1.735	2.088	2.700	3.325	16.919	19.023	21.666	23.589	9
10	2.156	2.558	3.247	3.940	18.307	20.483	23.209	25.188	10
11	2.603	3.053	3.816	4.575	19.675	21.920	24.725	26.757	11
12	3.074	3.571	4.404	5.226	21.026	23.337	26.217	28.300	12
13	3.565	4.107	5.009	5.892	22.362	24.736	27.688	29.819	13
14	4.075	4.660	5.629	6.571	23.685	26.119	29.141	31.319	14
15	4.601	5.229	6.262	7.261	24.996	27.488	30.578	32.801	15
16	5.142	5.812	6.908	7.962	26.296	28.845	32.000	34.267	16
17	5.697	6.408	7.564	8.672	27.587	30.191	33.409	35.718	17
18	6.265	7.015	8.231	9.390	28.869	31.526	34.805	37.156	18
19	6.844	7.633	8.907	10.117	30.144	32.852	36.191	38.582	19
20	7.434	8.260	9.591	10.851	31.410	34.170	37.566	39.997	20
21	8.034	8.897	10.283	11.591	32.671	35.479	38.932	41.401	21
22	8.643	9.542	10.982	12.338	33.924	36.781	40.289	42.796	22
23	9.260	10.196	11.689	13.091	35.172	38.076	41.638	44.181	23
24	9.886	10.856	12.401	13.848	36.415	39.364	42.980	45.558	24
25	10.520	11.524	13.120	14.611	37.652	40.646	44.314	46.928	25
26	11.160	12.198	13.844	15.379	38.885	41.923	45.642	48.290	26
27	11.808	12.879	14.573	16.151	40.113	43.194	46.963	49.645	27
28	12.461	13.565	15.308	16.928	41.337	44.461	48.278	50.993	28
29	13.121	14.256	16.047	17.708	42.557	45.722	49.588	52.336	29
30	13.787	14.953	16.791	18.493	43.773	46.979	50.892	53.672	30

The t distribution

Values of $t_{\alpha,\nu}$

ν	α = .10	α = .05	α = .025	α = .01	α = .005	ν
1	3.078	6.314	12.706	31.821	63.657	1
2	1.886	2.920	4.303	6.965	9.925	2
3	1.638	2.353	3.182	4.541	5.841	3
4	1.533	2.132	2.776	3.747	4.604	4
5	1.476	2.015	2.571	3.365	4.032	5
6	1.440	1.943	2.447	3.143	3.707	6
7	1.415	1.895	2.365	2.998	3.499	7
8	1.397	1.860	2.306	2.896	3.355	8
9	1.383	1.833	2.262	2.821	3.250	9
10	1.372	1.812	2.228	2.764	3.169	10
11	1.363	1.796	2.201	2.718	3.106	11
12	1.356	1.782	2.179	2.681	3.055	12
13	1.350	1.771	2.160	2.650	3.012	13
14	1.345	1.761	2.145	2.624	2.977	14
15	1.341	1.753	2.131	2.602	2.947	15
16	1.337	1.746	2.120	2.583	2.921	16
17	1.333	1.740	2.110	2.567	2.898	17
18	1.330	1.734	2.101	2.552	2.878	18
19	1.328	1.729	2.093	2.539	2.861	19
20	1.325	1.725	2.086	2.528	2.845	20
21	1.323	1.721	2.080	2.518	2.831	21
22	1.321	1.717	2.074	2.508	2.819	22
23	1.319	1.714	2.069	2.500	2.807	23
24	1.318	1.711	2.064	2.492	2.797	24
25	1.316	1.708	2.060	2.485	2.787	25
26	1.315	1.706	2.056	2.479	2.779	26
27	1.314	1.703	2.052	2.473	2.771	27
28	1.313	1.701	2.048	2.467	2.763	28
29	1.311	1.699	2.045	2.462	2.756	29
inf.	1.282	1.645	1.960	2.326	2.576	inf.

The F distribution

Values of $F_{.05,\nu_1,\nu_2}$

ν_1 = Degrees of freedom for numerator; ν_2 = Degrees of freedom for denominator

ν_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

Values of $F_{.01,\nu_1,\nu_2}$

ν_1 = Degrees of freedom for numerator; ν_2 = Degrees of freedom for denominator

ν_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4,052	5,000	5,403	5,625	5,764	5,859	5,928	5,982	6,023	6,056	6,106	6,157	6,209	6,235	6,261	6,287	6,313	6,339	6,366
2	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	99.4	99.5	99.5	99.5	99.5	99.5	99.5
3	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	27.1	26.9	26.7	26.6	26.5	26.4	26.3	26.2	26.1
4	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.4	14.2	14.0	13.9	13.8	13.7	13.7	13.6	13.5
5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.70	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.19	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.53	2.45	2.36	2.27	2.17
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

ORDER STATISTICS

Definition. The jth order statistic $\tilde{x}_{(j)}$ of a random sample $\{\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n\}$ of size n is the random variable of the sample that takes the jth smallest value.

Thus, the order statistics $\tilde{x}_{(1)}, \tilde{x}_{(2)}, ..., \tilde{x}_{(n)}$ of a random sample $\{\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n\}$ of size n are the same random variables of the sample relabelled in the same ascending order as the values they take. Thus, the values of the order statistics $\tilde{x}_{(1)}, \tilde{x}_{(2)}, ..., \tilde{x}_{(n)}$ satisfy $x_{(1)} \leq x_{(2)} \leq ... \leq x_{(n)}$.

We can define the values of the order statistics as

$$x_{(1)} = \min \{x_1, x_2, ..., x_n\},$$

$$x_{(2)} = \min \left\{ \{x_1, x_2, ..., x_n\} / \{x_{(1)}\} \right\},$$

$$\cdots$$

$$x_{(j)} = \min \left\{ \{x_1, x_2, ..., x_n\} / \{x_{(1)}, x_{(2)}, ..., x_{(j-1)}\} \right\},$$

$$\cdots$$

$$x_{(n)} = \min \left\{ \{x_1, x_2, ..., x_n\} / \{x_{(1)}, x_{(2)}, ..., x_{(n-1)}\} \right\} = \max \{x_1, x_2, ..., x_n\}.$$

There are some statistics that are easily defined in terms of the order statistics:

(1) The **minimum** $\tilde{x}_{\text{min}}$ of the random sample is $\tilde{x}_{(1)}$ and the **maximum** $\tilde{x}_{\text{max}}$ of the random sample is $\tilde{x}_{(n)}$, where n is the sample size.

(2) **The sample median**, which we will denote by $\tilde{x}_{\text{med}}$, is the number such that one half of the observations are smaller than $\tilde{x}_{\text{med}}$ and one-half are greater. Note that $\tilde{x}_{\text{med}}$ is random since each realization of the random sample will have its own realization of the median. In terms of order statistics, $\tilde{x}_{\text{med}}$ is defined by

$$\tilde{x}_{\text{med}} = \begin{cases} \tilde{x}_{\left(\frac{n+1}{2}\right)} & \text{if } n \text{ is odd} \\ \dfrac{\tilde{x}_{\left(\frac{n}{2}\right)} + \tilde{x}_{\left(\frac{n}{2}+1\right)}}{2} & \text{if } n \text{ is even.} \end{cases}$$

The median is a measure of location (or central tendency) that might be considered an alternative to the sample mean. One advantage of the sample median over the sample mean is that it is less affected by extreme observations (or outliers).

(3) **The sample range**, $\widetilde{R} = \tilde{x}_{(n)} - \tilde{x}_{(1)} = \tilde{x}_{\text{max}} - \tilde{x}_{\text{min}}$, is the distance between the largest and the smallest observations. It is a measure of the dispersion (or variability) in the sample and should reflect the dispersion in the population.

(4) **The p% sample percentile**. The value of the $p\%$ sample percentile, which we will denote by $x_{p\%}$ with $p \in [0, 100]$, is the value such that $pn/100$ of the observed values are smaller than $x_{p\%}$ and $(1-p)n/100$ of the observed values are

greater. In terms of order statistics, the statistic $\tilde{x}_{p\%}$ called $p\%$ sample percentile is defined by

$$\tilde{x}_{p\%} = \tilde{x}_{(j)},$$

where the integer j (between 1 and n) is the nearest to the value

$$\frac{pn}{100} + \frac{1}{2}. \tag{1}$$

If there were two integers j and $j+1$ that are at the same distance from (1), then the $p\%$ percentile is

$$\tilde{x}_{p\%} = \frac{\tilde{x}_{(j)} + \tilde{x}_{(j+1)}}{2}.$$

Note that the 50% percentile is the sample median, the 25% percentile is called the lower quartile, the 75% percentile is called the upper quartile, the 20% percentile is called the lower quintile, and the 80% percentile is called the upper quintile.

(5) **The interquartile range**, which is the distance between the lower and upper quartiles, $\tilde{x}_{75\%} - \tilde{x}_{25\%}$. This is also used as a measure of dispersion of the sample.

MEDIAN OF A POPULATION

The **median** m of a population $\tilde{x}$ having the distribution function F is the real number m for which $P\{\tilde{x} \leq m\} = F(m) \geq 1/2$ and $P\{\tilde{x} \geq m\} \geq 1/2$. If there are two values satisfying this definition, we take the average of them.

If the population $\tilde{x}$ has a continuous distribution function F, then the median m satisfies $F(m) = P\{\tilde{x} \leq m\} = 1/2$ (or, equivalently, $1 - F(m) = P\{\tilde{x} \geq m\} = 1/2$ since $P\{\tilde{x} = m\} = 0$).

The distribution of a random variable $\tilde{x}$ is symmetric with respect to (or around) $x^0 \in \mathbb{R}$ if $P\{\tilde{x} \leq x^0 - z\} = P\{\tilde{x} \geq x^0 + z\}$ for all $z \in \mathbb{R}$. Thus, a distribution is symmetric around x^0 if its probability function (density function) f satisfies $f(x^0 - z) = f(x^0 + z)$ for all z. Note that, when $\tilde{x} \in \mathrm{L}^3$, if the distribution is symmetric then its coefficient of skewness (or asymmetry) is equal to zero.

The **mode** of $\tilde{x}$ (or of its distribution) is the value (or the values) of the random variable for which its probability function (density function) f is at its global maximum. Sometimes, by modifying the density function on a set with zero Lebesgue measure, we can make

$$\max_x f(x) = \sup_x f(x).$$

A distribution is unimodal if it has a single mode.

If a random variable $\tilde{x}$ with well-defined mean μ (i.e., $\tilde{x} \in \mathrm{L}^1$) has a symmetric distribution, then the median and the mean are equal, $m = \mu$.

If a random variable $\tilde{x}$ with well-defined mean μ (i.e., $\tilde{x} \in \mathrm{L}^1$) has a symmetric and unimodal distribution, then the mode, the median, and the mean are all equal.

DISTRIBUTION OF THE MAXIMUM AND THE MINIMUM

Proposition 1. Let $\{\tilde{x}_i\}_{i=1}^n$ be a random sample of size n from a population $\tilde{x}$ having the distribution function (cdf) F. The distribution functions of the minimum $\tilde{x}_{\min} \equiv \tilde{x}_{(1)}$ and the maximum $\tilde{x}_{\max} \equiv \tilde{x}_{(n)}$ of the sample are

$$F_{\tilde{x}_{\min}}(x) = 1 - [1 - F(x)]^n \tag{2}$$

and

$$F_{\tilde{x}_{\max}}(x) = [F(x)]^n \,. \tag{3}$$

Proof. For the minimum $\tilde{x}_{\min}$,

$$\begin{aligned} F_{\tilde{x}_{\min}}(x) = F_{\tilde{x}_{(1)}}(x) &= P\left\{\tilde{x}_{(1)} \leq x\right\} = 1 - P\left\{\tilde{x}_{(1)} > x\right\} \\ &= 1 - P\left\{\tilde{x}_1 > x, \tilde{x}_2 > x, ..., \tilde{x}_n > x\right\} \\ &= 1 - P\left\{\tilde{x}_1 > x\right\} \cdot P\left\{\tilde{x}_2 > x\right\} \cdot ... \cdot P\left\{\tilde{x}_n > x\right\} \\ &= 1 - \left[P\left\{\tilde{x} > x\right\}\right]^n = 1 - \left[1 - P\left\{\tilde{x} \leq x\right\}\right]^n = 1 - [1 - F(x)]^n \,. \end{aligned}$$

For the maximum $\tilde{x}_{\max}$,

$$\begin{aligned} F_{\tilde{x}_{\max}}(x) = F_{\tilde{x}_{(n)}}(x) &= P\left\{\tilde{x}_{(n)} \leq x\right\} = P\left\{\tilde{x}_1 \leq x, \tilde{x}_2 \leq x, ..., \tilde{x}_n \leq x\right\} \\ &= P\left\{\tilde{x}_1 \leq x\right\} \cdot P\left\{\tilde{x}_2 \leq x\right\} \cdot ... \cdot P\left\{\tilde{x}_n \leq x\right\} \\ &\left[P\left\{\tilde{x} \leq x\right\}\right]^n = [F(x)]^n \,. \end{aligned}$$

Q.E.D.

Observe that

$$\lim_{n\to\infty} F_{\tilde{x}_{\min}}(x) = \lim_{n\to\infty} \left(1 - [1 - F(x)]^n\right) = \begin{cases} 0 & \text{if } F(x) = 0 \\ 1 & \text{if } F(x) > 0 \end{cases}$$

and

$$\lim_{n\to\infty} F_{\tilde{x}_{\max}}(x) = \lim_{n\to\infty} [F(x)]^n = \begin{cases} 0 & \text{if } F(x) < 1 \\ 1 & \text{if } F(x) = 1. \end{cases}$$

Let $M_F = \sup\left\{x \in \overline{\mathbb{R}} \,|F(x) < 1)\right\}$, then the extended random variable $\tilde{x}_{\max}$ converges in distribution as $n \to \infty$ to the constant M_F whether it is finite or infinite. Let $m_F = \inf\left\{x \in \overline{\mathbb{R}} \,|F(x) > 0)\right\}$, then the extended random variable $\tilde{x}_{\min}$ converges in distribution as $n \to \infty$ to the constant m_F whether it is finite or infinite. Thus, the distributions of both the minimum and the maximum of the random sample tend to be degenerate as the sample size goes to infinity.

DISTRIBUTION FUNCTION OF THE ORDER STATISTICS

Proposition 2. Let $\{\tilde{x}_i\}_{i=1}^n$ be a random sample of size n from a population $\tilde{x}$ having the distribution function F. Then, the distribution function $F_{\tilde{x}_{(j)}}$ of the jth order statistic $\tilde{x}_{(j)}$ is

$$F_{\tilde{x}_{(j)}}(x) = \sum_{k=j}^{n} \binom{n}{k} [F(x)]^k [1 - F(x)]^{n-k}. \tag{4}$$

Proof. Fix $x \in \mathbb{R}$, and let $\tilde{y}$ be a random variable that counts the number of random variables in $\{\tilde{x}_i\}_{i=1}^n$ that are smaller than or equal to x. For each of the random variables in $\{\tilde{x}_i\}_{i=1}^n$, call the event $\{\tilde{x}_i \leq x\}$ a "success" and $\{\tilde{x}_i > x\}$ a "failure". Then $\tilde{y}$ is the number of successes in n trials. Thus, $\tilde{y}$ is binomial with parameters n and $F(x)$, $\tilde{y} \sim \mathrm{B}(n, F(x))$. The event $\{\tilde{x}_{(j)} \leq x\}$ is equivalent to the event $\{\tilde{y} \geq j\}$, that is, at least j of the sample values are smaller than or equal to x. Then,

$$F_{\tilde{x}_{(j)}}(x) = P\left\{\tilde{x}_{(j)} \leq x\right\} = P\{\tilde{y} \geq j\} = \sum_{k=j}^{n} P\{\tilde{y} = k\} = \sum_{k=j}^{n} b(y; n, F(x)),$$

which becomes (4). *Q.E.D.*

Observe that equation (4) becomes (2) when $j = 1$. To see this, note that the binomial probability function $b(\cdot; n, F(x))$ satisfies

$$\sum_{k=0}^{n} b(x; n, F(x)) = \sum_{k=0}^{n} \binom{n}{k} [F(x)]^k [1 - F(x)]^{n-k} = 1$$

so that

$$\sum_{k=1}^{n} \binom{n}{k} [F(x)]^k [1 - F(x)]^{n-k} = 1 - \binom{n}{0} [F(x)]^0 [1 - F(x)]^{n-0} = 1 - [1 - F(x)]^n.$$

Therefore,

$$F_{\tilde{x}_{\min}}(x) = F_{\tilde{x}_{(1)}}(x) = \sum_{k=1}^{n} \binom{n}{k} [F(x)]^k [1 - F(x)]^{n-k} = 1 - [1 - F(x)]^n.$$

Moreover, when $j = n$, equation (4) becomes (3),

$$F_{\tilde{x}_{\max}}(x) = F_{\tilde{x}_{(n)}}(x) = \sum_{k=n}^{n} \binom{n}{k} [F(x)]^k [1 - F(x)]^{n-k}$$

$$= \binom{n}{n} [F(x)]^n [1 - F(x)]^0 = [F(x)]^n.$$

Proposition 3. Let $\{\tilde{x}_i\}_{i=1}^n$ be a random sample of size n from a discrete population $\tilde{x}$ having the probability function (pmf) $f : \tilde{x}(\Omega) \longrightarrow [0,1]$, where $\{x_1, x_2, ...\}$ are the different values of the range $\tilde{x}(\Omega)$ of $\tilde{x}$. Then, the probability function, $f_{\tilde{x}_{(j)}} : \tilde{x}(\Omega) \longrightarrow [0,1]$, of the jth order statistics $\tilde{x}_{(j)}$ is

$$f_{\tilde{x}_{(j)}}(x_i) =$$

$$\sum_{k=j}^{n} \binom{n}{k} \left(\underbrace{\left[\sum_{x \le x_i} f(x) \right]}_{P\{\tilde{x} \le x_i\}}^{k} \underbrace{\left[\sum_{x > x_i} f(x) \right]}_{P\{\tilde{x} > x_i\}}^{n-k} - \underbrace{\left[\sum_{x < x_i} f(x) \right]}_{P\{\tilde{x} < x_i\}}^{k} \underbrace{\left[\sum_{x \ge x_i} f(x) \right]}_{P\{\tilde{x} \ge x_i\}}^{n-k} \right)$$

for all $x_i \in \tilde{x}(\Omega)$.

Proof. It is immediate from using (4) and noticing that the probability function $f_{\tilde{x}_{(j)}}$ of $\tilde{x}_{(j)}$ satisfies

$$f_{\tilde{x}_{(j)}}(x_i) = F_{\tilde{x}_{(j)}}(x_i) - \lim_{x \to x_i^-} F_{\tilde{x}_{(j)}}(x), \quad \text{for all } x_i \in \tilde{x}(\Omega). \tag{5}$$

Therefore,

$$f_{\tilde{x}_{(j)}}(x_i) =$$

$$\sum_{k=j}^{n} \binom{n}{k} \left([F(x_i)]^k [1 - F(x_i)]^{n-k} - \left[\lim_{x \to x_i^-} F(x) \right]^k \left[1 - \lim_{x \to x_i^-} F(x) \right]^{n-k} \right),$$

for all $x_i \in \tilde{x}(\Omega)$. Since $F(x_i) = \sum_{x \le x_i} f(x) = P\{\tilde{x} \le x_i\}$ and $\lim_{x \to x_i^-} F(x) = \sum_{x < x_i} f(x) = P\{\tilde{x} < x_i\}$, we get the desired result. *Q.E.D.*

The previous proposition implies that the value of the probability function of the jth order statistics evaluated at the lowest value, say x_1, of the range $\tilde{x}(\Omega)$ of the discrete population $\tilde{x}$ is

$$f_{\tilde{x}_{(j)}}(x_1) = F_{\tilde{x}_{(j)}}(x_1) = \sum_{k=j}^{n} \binom{n}{k} [F(x_1)]^k [1 - F(x_1)]^{n-k}$$

$$= \sum_{k=j}^{n} \binom{n}{k} [f(x_1)]^k [1 - f(x_1)]^{n-k}$$

since $\lim_{x \to x_1^-} F(x) = 0$ and $F(x_1) = f(x_1)$.

Corollary 1. Let $\{\tilde{x}_i\}_{i=1}^n$ be a random sample of size n from a discrete population $\tilde{x}$ having the probability function (pmf) $f : \tilde{x}(\Omega) \longrightarrow [0,1]$, where $\{x_1, x_2, ...\}$

are the different values of the range $\tilde{x}(\Omega)$ of $\tilde{x}$. Then, the probability functions $f_{\tilde{x}_{\min}}$ and $f_{\tilde{x}_{\max}}$ of the minimum and of the maximum of the sample are

$$f_{\tilde{x}_{\min}}(x_i) = \underbrace{\left[\sum_{x \geq x_i} f(x)\right]^n}_{P\{\tilde{x} \geq x_i\}} - \underbrace{\left[\sum_{x > x_i} f(x)\right]^n}_{P\{\tilde{x} > x_i\}} \text{ for all } x_i \in \tilde{x}(\Omega),$$

and

$$f_{\tilde{x}_{\max}}(x_i) = \underbrace{\left[\sum_{x \leq x_i} f(x)\right]^n}_{P\{\tilde{x} \leq x_i\}} - \underbrace{\left[\sum_{x < x_i} f(x)\right]^n}_{P\{\tilde{x} < x_i\}} \text{ for all } x_i \in \tilde{x}(\Omega).$$

Proof. Combine (5) for $j = 1$ and for $j = n$ with (2) and (3), respectively, to get

$$f_{\tilde{x}_{\min}}(x_i) = \left[1 - \lim_{x \to x_i^-} F(x)\right]^n - [1 - F(x_i)]^n$$

and

$$f_{\tilde{x}_{\max}}(x_i) = [F(x_i)]^n - \left[\lim_{x \to x_i^-} F(x)\right]^n,$$

for all $x_i \in \tilde{x}(\Omega)$, and the result immediately follows. *Q.E.D.*

DENSITY FUNCTION OF THE ORDER STATISTICS

Proposition 4. Let $\{\tilde{x}_i\}_{i=1}^n$ be a random sample of size n from an absolutely continuous population $\tilde{x}$ having the density function (pdf) f and the distribution function F. Then, the density $f_{\tilde{x}_{(j)}}$ of the jth order statistic is

$$f_{\tilde{x}_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} f(x) [F(x)]^{j-1} [1 - F(x)]^{n-j}. \tag{6}$$

Proof. One way of proving this proposition is to compute the derivative of (4) with respect to x. This is a little bit painful.

We will use an alternative approach in our proof. We can assume safely that $x_{(1)} < x_{(2)} < ... < x_{(n)}$ since $P\{\tilde{x}_i = \tilde{x}_j\} = 0$ for all $i \neq j$ when the population is continuous. Then, we divide the real axis into the following three intervals: $(-\infty, x)$, $[x, x+h]$, and $(x+h, \infty)$. Then, the probability that the jth order statistic $\tilde{x}_{(j)}$ falls into the interval $[x, x+h]$ is equal to the probability that $j-1$ of the sample values fall into the interval $(-\infty, x)$, one falls into the interval $[x, x+h]$, and $n-j$ fall into the interval $(x+h, \infty)$. Therefore, since the distribution function F of $\tilde{x}$ is continuous, we have

$$P\left\{\tilde{x}_{(j)} \in [x, x+h]\right\} =$$

$$\frac{n!}{(j-1)!1!(n-j)!}\left[P\left\{\tilde{x} \in (-\infty, x)\right\}\right]^{j-1} \cdot P\left\{\tilde{x} \in [x, x+h]\right\} \cdot \left[P\left\{\tilde{x} \in (x+h, \infty)\right\}\right]^{n-j}$$

$$= \frac{n!}{(j-1)!(n-j)!}\left[F(x)\right]^{j-1} \cdot P\left\{\tilde{x} \in [x, x+h]\right\} \cdot \left[1 - F(x+h)\right]^{n-j}, \qquad (7)$$

according to the formula of the multinomial distribution.

We know that, for an absolutely continuous random variable $\tilde{y}$,

$$P\left\{\tilde{y} \in [x, x+h]\right\} = F_{\tilde{y}}(x+h) - F_{\tilde{y}}(x),$$

where $F_{\tilde{y}}$ is the (continuous) distribution function of $\tilde{y}$. Moreover, using the mean value theorem, there exists a value $\varepsilon_{\tilde{y}} \in [x, x+h]$ such that

$$F_{\tilde{y}}(x+h) - F_{\tilde{y}}(x) = F'_{\tilde{y}}(\varepsilon_{\tilde{y}})h = f_{\tilde{y}}(\varepsilon_{\tilde{y}})h, \quad \text{a.e.},$$

where $f_{\tilde{y}}$ is the density of $\tilde{y}$.

Thus, (7) becomes

$$f_{\tilde{x}_{(j)}}(\varepsilon_{\tilde{x}_{(j)}})h = \frac{n!}{(j-1)!(n-j)!}\left[F(x)\right]^{j-1} \cdot f(\varepsilon_{\tilde{x}})h \cdot \left[1 - F(x+h)\right]^{n-j},$$

for some $\varepsilon_{\tilde{x}_{(j)}} \in [x, x+h]$ and $\varepsilon_{\tilde{x}} \in [x, x+h]$. If we let $h \to 0$ and divide by h, the previous expression becomes

$$f_{\tilde{x}_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!}\left[F(x)\right]^{j-1} \cdot f(x) \cdot \left[1 - F(x)\right]^{n-j}. \qquad \textit{Q.E.D.}$$

Corollary 2. Let $\{\tilde{x}_i\}_{i=1}^n$ be a random sample of size n from an absolutely continuous population $\tilde{x}$ having the density f and the distribution function F. The density functions $f_{\tilde{x}_{\min}}$ and $f_{\tilde{x}_{\max}}$ of the minimum and the maximum of the sample are

$$f_{\tilde{x}_{\min}}(x) = nf(x)\left[1 - F(x)\right]^{n-1} \qquad (8)$$

and

$$f_{\tilde{x}_{\max}}(x) = nf(x)\left[F(x)\right]^{n-1}. \qquad (9)$$

Proof. Just compute the derivative with respect to x of the distribution functions obtained in Proposition 1 or, alternatively, evaluate (6) at $j = 1$ and at $j = n$. *Q.E.D.*

DENSITY FUNCTION OF THE MEDIAN

Consider a random sample $\{\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n\}$ of size n, where n is an odd number. Obviously, the sample median $\tilde{x}_{\text{med}}$ is the $((n+1)/2)$th order statistic. Then, using equation (6) for this statistic, we get immediately the following:

Proposition 4. Let $\{\tilde{x}_i\}_{i=1}^n$ be a random sample of size n (with n being an odd number) from an absolutely continuous population $\tilde{x}$ having the density f and the distribution function F. Then, the density $f_{\tilde{x}_{\text{med}}}$ of the sample median $\tilde{x}_{\text{med}}$ is

$$f_{\tilde{x}_{\text{med}}}(x) = \frac{n!}{\left[\left(\frac{n-1}{2}\right)!\right]^2} f(x) \left[F(x)\right]^{(n-1)/2} \left[1 - F(x)\right]^{(n-1)/2}.$$

We next provide with a nice proposition, which we state without a proof.

Proposition 5. Let $\{\tilde{x}_i\}_{i=1}^n$ be a random sample of size n from an absolutely continuous population $\tilde{x}$ having the median m and the density f. Assume that $f(m) > 0$ and that f is continuously differentiable in a neighborhood of m. Then, as $n \to \infty$,

$$2\sqrt{n} f(m) \left(\tilde{x}_{\text{med}} - m\right) \longrightarrow \mathrm{N}(0,1), \tag{10}$$

or, equivalently,

$$\sqrt{n}\left(\tilde{x}_{\text{med}} - m\right) \longrightarrow \mathrm{N}\left(0, \frac{1}{4\left[f(m)\right]^2}\right). \tag{11}$$

Corollary 3. Let $\{\tilde{x}_i\}_{i=1}^n$ be a random sample of size n from a normal population $\tilde{x}$ with the mean μ and the variance σ^2, $\tilde{x} \sim \mathrm{N}(\mu, \sigma^2)$. Then, as $n \to \infty$,

$$\sqrt{\frac{2}{\pi}} \left(\frac{\tilde{x}_{\text{med}} - \mu}{\sigma/\sqrt{n}}\right) \longrightarrow \mathrm{N}(0,1), \tag{12}$$

or, equivalently,

$$\sqrt{n}\left(\tilde{x}_{\text{med}} - \mu\right) \longrightarrow \mathrm{N}\left(0, \frac{\pi\sigma^2}{2}\right). \tag{13}$$

Proof. First, note that $m = \mu$ since the normal distribution is symmetric. Moreover, the normal density evaluated at μ is equal to $f(\mu) = 1/\sigma\sqrt{2\pi}$. Therefore, the expression (10) becomes (12), while (11) becomes (13). *Q.E.D.*

If the population is normal, $\tilde{x} \sim \mathrm{N}(\mu, \sigma^2)$, the previous corollary tells us that, for n large, $\mathrm{E}(\tilde{x}_{\text{med}}) \approx \mu$. Remember that the sample mean $\bar{\boldsymbol{x}}$, satisfies $\mathrm{E}(\bar{\boldsymbol{x}}) = \mu$. Thus, for n large, both $\tilde{x}_{\text{med}}$ and $\bar{\boldsymbol{x}}$ have approximately the same mean, which is the population mean μ. However, for n large, the sample median $\tilde{x}_{\text{med}}$ has a larger variance than the sample mean $\bar{\boldsymbol{x}}$,

$$\mathrm{Var}(\tilde{x}_{\text{med}}) \approx \frac{\pi\sigma^2}{2n} > \mathrm{E}(\bar{\boldsymbol{x}}) = \frac{\sigma^2}{n},$$

as follows from (13) and the fact that $\pi/2 > 1$.

EXTREME VALUE THEORY

We have already seen that, if $\{\tilde{x}_1, \tilde{x}_2, ...\}$ are i.i.d. random variables with common cdf F, then $\tilde{x}^n_{\max} = \max\{\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n\}$ converges in distribution as $n \to \infty$ to the constant $M(F) = \sup\{x \in \overline{\mathbb{R}} | F(x) < 1)\}$ whether it is finite or infinite. Thus, the distribution of the maximum of the random sample tends to be degenerate as the sample size goes to infinity.

Extremal types theorem. Assume that for suitable constants $c_n > 0$ and b_n, the random variable $\tilde{z}_n = \dfrac{\tilde{x}^n_{\max} - b_n}{c_n}$ converges in distribution to a non-degenerate random variable, $F_{\tilde{z}_n} \xrightarrow{w} G$. Then, the limiting distribution function G is of one of the following three classes:
(i)

$$G(x) = e^{-e^{-\left(\frac{x-m}{s}\right)}}, \quad \text{for all } x \in \mathbb{R},$$

for some $m \in \mathbb{R}$ and $s > 0$.
(ii)

$$G(x) = \begin{cases} 0 & \text{for } x < m \\ e^{-\left(\frac{x-m}{s}\right)^{-\alpha}} & \text{for } x \geq m. \end{cases}$$

for some $m \in \mathbb{R}$, $s > 0$, and $\alpha > 0$.
(iii)

$$G(x) = \begin{cases} e^{-\left(-\frac{x-m}{s}\right)^{\alpha}} & \text{for } x < m \\ 1 & \text{for } x \geq m \end{cases}$$

for some $m \in \mathbb{R}$, $s > 0$, and $\alpha > 0$.

Note that the classes (i) and (ii) correspond to the generalized extreme value type I (Gumbel) and type II (Fréchet) distributions, respectively. Concerning the class III, recall that the generalized extreme value type III (Weibull) distribution has the following distribution function:

$$F(y; m_0, s, \alpha) = P(\tilde{y} \leq y) = \begin{cases} 0 & \text{for } y < m_0 \\ 1 - e^{-\left(\frac{y-m_0}{s}\right)^{\alpha}} & \text{for } y \geq m_0. \end{cases}$$

for $m \in \mathbb{R}$, $s > 0$, and $\alpha > 0$. Therefore, the distribution function of the random variable $\tilde{x} = -\tilde{y}$ will be

$$F_{\tilde{x}}(x) = P(\tilde{x} \leq x) = P(-\tilde{y} \leq x) = P(\tilde{y} \geq -x) = 1 - P(\tilde{y} \leq -x) = 1 - F(-x; m_0, s, \alpha)$$

so that

$$F_{\tilde{x}}(x) = 1 - F(-x; m_0, s, \alpha) = \begin{cases} 1 - 0 = 1 & \text{for } -x < m_0 \text{ or } x \geq -m_0 \\ 1 - \left[1 - e^{-\left(\frac{-x-m_0}{s}\right)^{\alpha}}\right] = e^{-\left(\frac{-x-m_0}{s}\right)^{\alpha}} & \text{for } -x \geq m_0 \text{ or } x < -m_0. \end{cases}$$

Making $m = -m_0 \in \mathbb{R}$, the previous distribution function becomes

$$F_{\tilde{x}}(x) = \begin{cases} 1 & \text{for } x \geq m \\ \\ e^{-\left(\frac{-x+m}{s}\right)^{\alpha}} = e^{-\left(-\frac{x-m}{s}\right)^{\alpha}} & \text{for } x < m, \end{cases}$$

which coincides with the distribution function of the class (iii). Thus, the distribution of the class (iii) is the distribution of the negative of a random variable that has the generalized extreme value type III (or Weibull) distribution. This distribution is called **negative Weibull.**

Exercises. Probability and Statistics. IDEA.
7. Sampling

1. Verify the following computing formula for the value of the sample variance:

$$s^2 = \frac{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}{n\,(n-1)}.$$

2. If $\tilde{x}_1, \tilde{x}_2, ...,$ and $\tilde{x}_n$ are independent random variables having identical Bernoulli distributions with the parameter θ, then the sample mean $\bar{\boldsymbol{x}}_{\mathbf{n}}$ is the proportion of successes in n trials.

   (a) Verify that $\mathrm{E}(\bar{\boldsymbol{x}}_{\mathbf{n}}) = \theta$ and $\mathrm{Var}(\bar{\boldsymbol{x}}_{\mathbf{n}}) = \dfrac{\theta\,(1-\theta)}{n}$.

   (b) Using the moment generating function method, prove that, if $\tilde{x}_1, ..., \tilde{x}_n$ are independently distributed random variables having a Bernoulli distribution with parameter θ, then $\tilde{y}_n = \sum_{i=1}^{n} \tilde{x}_i$ has a binomial distribution with parameters n and θ.

   (c) Use parts (a) and (b), and the central limit theorem to prove that

$$\frac{\tilde{y}_n - n\theta}{\sqrt{n\theta\,(1-\theta)}} \longrightarrow \mathrm{N}\,(0,1)$$

3. A random sample of size $n = 100$ is taken from a population with the mean $\mu = 75$ and the variance $\sigma^2 = 256$. Use Chebyshev's inequality to provide a lower bound for the probability that the value of the sample mean $\bar{\boldsymbol{x}}$ falls between 67 and 83?

4. Use the central limit theorem to find an approximate value for the probability of Exercise 3.

5. A random sample of size 100 is taken from a normal population with $\sigma = 25$. Find the probability that the mean of the sample will differ from the mean of the population by 3 or more either way.

6. Find approximate values for the probability that a random variable $\tilde{x}$ having a chi-square distribution with 50 degrees of freedom will take on a value greater than 68.

   (a) by treating $\dfrac{\tilde{x}-\nu}{\sqrt{2\nu}}$ with $\nu = 50$ as a random variable having the standard normal distribution;

   (b) by treating $\sqrt{2\tilde{x}} - \sqrt{2\nu}$ with $\nu = 50$ as a random variable having the standard normal distribution.

   Also, judge the merits of these approximations, given that the actual value of the probability (rounded to five decimals) is 0.04596.

   (c) Justify the approximations given in (a) and (b). *Hint:* You should prove that, for all ν and for every positive real valued random variable $\tilde{x}$, the probability that the random variable $\sqrt{2\tilde{x}} - \sqrt{2\nu}$ takes on a value less than k equals the probability that the random variable $\dfrac{\tilde{x} - \nu}{\sqrt{2\nu}}$ takes on a value less than $k + \dfrac{k^2}{2\sqrt{2\nu}}$.

7. Show that the density of a F distribution with 4 and 4 degrees of freedom is given by

$$g(y) = \begin{cases} 6y\,(1+y)^{-4} & \text{for } y > 0 \\ 0 & \text{elsewhere} \end{cases}$$

and use this density to find the probability that for independent random samples of size 5 from normal populations having the same variance, $\dfrac{\mathbf{s_1^2}}{\mathbf{s_2^2}}$ will take on a value less than 1/2 or greater than 2.

8. If $\tilde{x}$ has an F distribution with ν_1 and ν_2 degrees of freedom, show that $\tilde{y} = \dfrac{1}{\tilde{x}}$ has the F distribution with ν_2 and ν_1 degrees of freedom.

9. If $\tilde{x}$ has an F distribution with ν_1 and ν_2 degrees of freedom, show that when $\nu_2 \to \infty$ the distribution of $\nu_1\tilde{x}$ approaches the chi-square distribution with ν_1 degrees of freedom.

10. Show that if $\tilde{x}$ has the t distribution with ν degrees of freedom, then $\tilde{x}^2$ has the F distribution with 1 and ν degrees of freedom.

11. The claim that the variance of a normal population is $\sigma^2 = 25$ is to be rejected if the variance of a random sample of size 16 exceeds 54.668 or is less than 12.102. What is the probability that this claim will be rejected even though $\sigma^2 = 25$?

12. Let $\boldsymbol{s}_n$ be the standard deviation of a random sample of size n from a normal population. Use the approximation given in part (b) of Exercise 6, to prove that for large n the variance of the sampling distribution of $\boldsymbol{s}_n$ is approximately $\dfrac{\sigma^2}{2(n-1)}$.

13. Consider an experiment consisting of randomly selecting n values from a finite set of distinct numbers $C = \{c_1, c_2, ..., c_N\}$ with $N \geq n$. Let us assume that the selection is without replacement and $\tilde{x}_1$ is the first number drawn, $\tilde{x}_2$ is the second number drawn,..., and $\tilde{x}_n$ is the nth number drawn.

   (a) Find the joint probability function $f(x_1, x_2, ..., x_n)$ for each ordered n-tuple of distinct values selected from the set C.

   (b) Find the probability of each subset of n of the N elements of the set C, regardless of the order in which the values are obtained.

   (c) Find the marginal probability function of each $\tilde{x}_i$, $f(x_i)$ for $x_i = c_1, c_2, ..., c_N$.

   (d) Find the mean μ and the variance σ^2 of each $\tilde{x}_i$.

(e) Find the marginal probability function of any two of the random variables $\{\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n\}$, $g(x_i, x_j)$, for each ordered pair (x_i, x_j) of distinct values of the set C.

For the rest of this exercise let us assume that we do not know the values $\{c_1, c_2, ..., c_N\}$, but we know that the mean and the variance of each $\tilde{x}_i$ are μ and σ^2, respectively.

(f) Find the covariance of any two of the random variables $\{\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n\}$, $\mathrm{Cov}(\tilde{x}_i, \tilde{x}_j)$ with $i \neq j$. *Note:* Your answer should be a function of σ^2 and N only since $\mathrm{Cov}(\tilde{x}_i, \tilde{x}_j)$ turns out to be independent of n and μ.

(g) Let $\bar{\boldsymbol{x}}$ be the average of the random variables $\{\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n\}$, $\bar{\boldsymbol{x}} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} \tilde{x}_i$.

Find $\mathrm{Var}(\bar{\boldsymbol{x}})$. *Note:* Your answer should be a function of σ^2, N and n only since $\mathrm{Var}(\bar{\boldsymbol{x}})$ turns out to be independent of μ.

(h) Compute $\underset{N\to\infty}{\mathrm{Var}}(\bar{\boldsymbol{x}})$. Discuss this result.

14. The density function of the random vector $(\tilde{x}, \tilde{y})$ is

$$f(x,y) = \begin{cases} \dfrac{1}{\pi} e^{-\frac{1}{2}(x^2+y^2)} & \text{for } 0 < x < \infty, \quad -\infty < y < \infty \\ \\ 0 & \text{otherwise.} \end{cases}$$

(a) Prove that $f(x,y)$ is the joint density of two independent random variables $\tilde{x}$ and $\tilde{y}$, where $\tilde{x}$ is the absolute value of a standard normal random variable and $\tilde{y}$ is standard normal.

(b) Find $\mathrm{E}(\tilde{x})$, $\mathrm{E}(\tilde{y})$, $\mathrm{Var}(\tilde{x})$, $\mathrm{Var}(\tilde{y})$, and $\mathrm{Cov}(\tilde{x}, \tilde{y})$.

(c) Find the density function of the random variable $\tilde{v} = \dfrac{\tilde{y}}{\tilde{x}}$. Show all the computations.

(d) Prove that the random variable $\tilde{v}$ has a Student's t distribution with one degree of freedom, $\tilde{v} \sim \mathrm{t}_1$.

15. Assume that $\bar{\boldsymbol{x}}$ is the mean of a random sample $\{\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n\}$ of size n from a population having the mean μ and the finite variance σ^2. Note that $\bar{\boldsymbol{x}}$ is a random variable.

Find

(a) $\mathrm{Cov}(\tilde{x}_i, \tilde{x}_j)$, for $i \neq j$.

(b) $\mathrm{Cov}(\bar{\boldsymbol{x}}, \tilde{x}_i)$, for $i = 1, ..., n$.

(c) $\mathrm{Cov}(\tilde{x}_i, \tilde{x}_i - \bar{\boldsymbol{x}})$, for $i = 1, ..., n$.

(d) $\mathrm{Cov}(\tilde{x}_i, \tilde{x}_j - \bar{\boldsymbol{x}})$, for $i \neq j$.

(e) $\mathrm{Var}(\tilde{x}_i - \bar{\boldsymbol{x}})$, for $i = 1, ..., n$.

(f) $\mathrm{Cov}(\tilde{x}_i - \bar{\boldsymbol{x}}, \tilde{x}_j - \bar{\boldsymbol{x}})$, for $i \neq j$.

(g) $\mathrm{Cov}\left(\bar{\boldsymbol{x}}, \tilde{x}_i - \bar{\boldsymbol{x}}\right)$, for $i = 1, ..., n$.

Hint: You can use the results of Exercise 29 of List 3.

16. (a) Prove that, as the number ν of degrees of freedom goes to infinity, the limit of the t density is the standard normal density.

(b) Prove that the t distribution with one degree of freedom is a Cauchy distribution with parameters $\alpha = 0$ and $\beta = 1$. Recall that the density of the Cauchy distribution with parameters α and β is

$$f(x) = \frac{\beta/\pi}{(x-\alpha)^2 + \beta^2} \quad \text{for} - \infty < x < \infty.$$

(c) Assume that the random variable $\tilde{x} : (\Omega, \mathcal{F}, P) \longrightarrow (\mathbb{R}, \mathcal{B})$ has the t distribution with one degree of freedom. Compute the conditional expectation of $\tilde{x}$ given that $\tilde{x} \in [0, 1]$. Recall that

$$\mathrm{E}\left(\tilde{x} \mid \tilde{x} \in [0,1]\right) \equiv \mathrm{E}\left(\tilde{x} \mid \mathbb{I}_A = 1\right),$$

where $\mathbb{I}_A$ is the indicator function of the set $A = \{\omega \in \Omega \mid \tilde{x}(\omega) \in [0,1]\}$. Moreover, if $\psi(x) = \arctan(x)$, where the function $\arctan(\cdot) : \mathbb{R} \longrightarrow (-\pi/2, \pi/2)$ is the inverse function of $\tan(\cdot) : (-\pi/2, \pi/2) \longrightarrow \mathbb{R}$, then the derivative of ψ is $\psi'(x) = \dfrac{1}{1+x^2}$. Recall also that $\tan\left(\dfrac{\pi}{4}\right) = 1$, $\tan\left(-\dfrac{\pi}{4}\right) = -1$, $\lim\limits_{\theta \to \pi/2} \tan(\theta) = \infty$, $\lim\limits_{\theta \to -\pi/2} \tan(\theta) = -\infty$ and $\tan(0) = 0$.

17. Let $\{\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n\}$ be a random sample of size n from a population $\tilde{x}$ having the uniform density on $(0,1)$.

(a) Find the density $f_{\tilde{x}_{(j)}}$ of the jth order statistic. Could you identify the distribution of the jth order statistic $\tilde{x}_{(j)}$? Find the expectation of $\tilde{x}_{(j)}$. Plot the density of the 2nd order statistic $\tilde{x}_{(2)}$ when the sample size is $n = 7$ and find the expectation of $\tilde{x}_{(2)}$ in this case.

(b) Find the expectation of the sample median $\tilde{x}_{\mathrm{med}}$. Find the density of the sample median when the sample size is $n = 7$. Plot this density.

18. Let $\{\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n\}$ be a random sample of size n from a population $\tilde{x}$ having the exponential density with the parameter θ.

(a) Find the densities, $f_{\tilde{x}_{\min}}$ and $f_{\tilde{x}_{\max}}$, and the distribution functions, $F_{\tilde{x}_{\min}}$ and $F_{\tilde{x}_{\max}}$, of the minimum and the maximum of the sample, respectively. Could you identify the distribution of the sample minimum $\tilde{x}_{\min}$? Find the expectation of $\tilde{x}_{\min}$. Plot the density of the sample maximum $\tilde{x}_{\max}$ when $\theta = 3$ and $n = 11$ and find the expectation of both $\tilde{x}_{\min}$ and $\tilde{x}_{\max}$ in this case.

(b) Find the median and the mode of the population distribution and compare them with the population mean. Find the density of the sample median $\tilde{x}_{\mathrm{med}}$ when n is an odd number. Plot the density of $\tilde{x}_{\mathrm{med}}$ when $\theta = 3$ and $n = 11$. For this case, find the expectation of $\tilde{x}_{\mathrm{med}}$ and compare it with the expectation of the sample mean, the population mean, the population median, the mode of the population distribution and the mode of the distribution of $\tilde{x}_{\text{med}}$.

19. Let $\{\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n\}$ be a collection of independent random variables having the Poisson distribution with the parameters λ_i, for $i = 1, 2, ..., n$.

    (a) Use the formula for the convolution pmf to prove that the sum $\tilde{S} = \sum_{i=1}^{n} \tilde{x}_i$ has a Poisson distribution with the parameter $\sum_{i=1}^{n} \lambda_i$.

    (b) Find the probability function of the average $\bar{\boldsymbol{x}} = \dfrac{\sum_{i=1}^{n} \tilde{x}_i}{n}$.

    (c) Assume now that $\lambda_i = \lambda$, for $i = 1, 2, ..., n$, which implies that $\{\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n\}$ is a random sample. Find the probability function of the sample mean $\bar{\boldsymbol{x}}$.

    (d) Under the assumption of part (c) prove that $\sqrt{n}\,(\bar{\boldsymbol{x}}_n - \lambda) \longrightarrow \mathrm{N}(0, \lambda)$, as $n \longrightarrow \infty$, i.e., $\sqrt{n}\,(\bar{\boldsymbol{x}}_n - \lambda)$ is asymptotically distributed as a normal random variable with zero mean and variance λ, where $\bar{\boldsymbol{x}}_{\boldsymbol{n}} = \dfrac{\sum_{i=1}^{n} \tilde{x}_i}{n}$.

20. Consider a random sample of size n from an absolutely continuous population. Compute the derivative of the distribution function $F_{\tilde{x}_{(j)}}(x)$ of the jth order statistic given in the expression (4) of the handout "Order Statistics" to obtain the corresponding density $f_{\tilde{x}_{(j)}}(x)$, which is given in the expression (6).

21. We say that a random variable $\tilde{x}$ has a distribution symmetric with respect to x^0 if $P\{\tilde{x} \leq x^0 - z\} = P\{\tilde{x} \geq x^0 + z\}$ for all $z \in \mathbb{R}$.

    (a) Prove that $\tilde{x}$ has a distribution symmetric with respect to x^0 if and only if the random variable $\tilde{y} = \tilde{x} - x^0$ has a distribution symmetric with respect to zero.

    (b) Prove that the random variable $\tilde{y}$ has a distribution symmetric with respect to zero if and only if $\tilde{y}$ has the same distribution as the random variable $-\tilde{y}$.

    (c) Assume that the random variable $\tilde{y}$ is discrete (absolutely continuous). Prove that $\tilde{y}$ has a distribution symmetric with respect to zero if and only if its probability (density) function satisfies $f_{\tilde{y}}(-y) = f_{\tilde{y}}(y)$ for all y, i.e., the function $f_{\tilde{y}}$ is symmetric with respect to zero.

    (d) Assume that the random variable $\tilde{x}$ is discrete (absolutely continuous). Prove that $\tilde{x}$ has a distribution symmetric with respect to x^0 if and only if its probability (density) function satisfies $f_{\tilde{x}}(x^0 - z) = f_{\tilde{x}}(x^0 + z)$ for all z, i.e., the function $f_{\tilde{x}}$ is symmetric with respect to x^0.

    For the rest of the exercise assume that the random variable $\tilde{x}$ has a distribution symmetric with respect to x^0.

    (e) Prove that $\mathrm{E}(\tilde{x}) = x^0$ when $\tilde{x} \in \mathrm{L}^1$.

    (f) Prove that the coefficient of asymmetry (or skewness) of $\tilde{x}$ is equal to zero when $\tilde{x} \in \mathrm{L}^3$.