

Sampling Distributions

UGC NET Statistics | Unit: Sampling Theory | Concept + Formulas + Solved Examples

1. Definitions

SAMPLING DISTRIBUTION

If $X_1, \dots, X_n$ is a random sample from a population, and $T = T(X_1, \dots, X_n)$ is a statistic (e.g. sample mean $\bar{X}$), the probability distribution of T across all possible samples is called its **sampling distribution**.

STANDARD ERROR

The **standard error** of a statistic T is its standard deviation: $SE(T) = \sqrt{\text{Var}(T)}$. For the sample mean: $SE(\bar{X}) = \frac{\sigma}{\sqrt{n}}$.

2. Central Limit Theorem

THEOREM — CENTRAL LIMIT THEOREM (CLT)

If $X_1, \dots, X_n$ are i.i.d. with mean μ and finite variance σ^2 , then as $n \rightarrow \infty$:

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0, 1)$$

This holds **regardless of the population's original distribution** — the key reason normal-based inference works so broadly.

3. Chi-Square, t, and F Distributions

CHI-SQUARE DISTRIBUTION

If $Z_1, \dots, Z_n$ are i.i.d. $N(0, 1)$, then:

$$\sum_{i=1}^n Z_i^2 \sim \chi_n^2$$

with n **degrees of freedom**. Mean = n , Variance = $2n$. Sum of independent chi-squares is chi-square: $\chi_m^2 + \chi_n^2 \sim \chi_{m+n}^2$.

STUDENT'S T -DISTRIBUTION

If $Z \sim N(0, 1)$ and $W \sim \chi_n^2$ are independent:

$$T = \frac{Z}{\sqrt{W/n}} \sim t_n$$

t_n is symmetric about 0, has heavier tails than $N(0, 1)$, and $t_n \rightarrow N(0, 1)$ as $n \rightarrow \infty$.

F -DISTRIBUTION

If $U \sim \chi_m^2$ and $V \sim \chi_n^2$ are independent:

$$F = \frac{U/m}{V/n} \sim F_{m,n}$$

Key relation: if $T \sim t_n$, then $T^2 \sim F_{1,n}$.

4. Key Sampling Results for Normal Populations

SAMPLE MEAN AND VARIANCE (NORMAL POPULATION)

If $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ i.i.d., with $\bar{X} = \frac{1}{n} \sum X_i$ and $S^2 = \frac{1}{n-1} \sum (X_i - \bar{X})^2$:

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right), \quad \frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$$

and crucially, $\bar{X}$ and S^2 are **independent** for normal populations (a special property unique to the normal distribution).

T-STATISTIC FOR UNKNOWN VARIANCE

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

5. Key Points to Remember

- χ^2 , t , and F distributions are all **derived from the normal distribution** — they exist to handle inference when the population variance is unknown or when comparing variances.
- As degrees of freedom $\rightarrow \infty$: $t_n \rightarrow N(0, 1)$, and $\chi_n^2/n \rightarrow 1$ (in probability).
- $F_{m,n}$ and $F_{n,m}$ are reciprocally related: if $F \sim F_{m,n}$, then $1/F \sim F_{n,m}$.
- The independence of $\bar{X}$ and S^2 is a **special property of the normal distribution** — it does not hold for arbitrary distributions.
- CLT justifies using normal approximations for sample means from any population, given sufficiently large n (commonly $n \geq 30$ as a rule of thumb).
- Degrees of freedom for S^2 is $n - 1$, not n , because one degree of freedom is used in estimating μ via $\bar{X}$.

6. Worked Examples

EXAMPLE 1 — APPLYING CLT

A population has mean $\mu = 50$, $\sigma = 12$. For $n = 36$, find $P(\bar{X} > 53)$.

Solution: $SE(\bar{X}) = 12/\sqrt{36} = 2$. By CLT, $\bar{X} \approx N(50, 4)$.

$$P(\bar{X} > 53) = P\left(Z > \frac{53 - 50}{2}\right) = P(Z > 1.5) \approx 0.0668$$

EXAMPLE 2 — CHI-SQUARE SUM

If $X \sim \chi_5^2$ and $Y \sim \chi_8^2$ are independent, find the distribution of $X + Y$ and its mean.

Solution: $X + Y \sim \chi_{13}^2$ (df add). Mean = 13 (since mean of χ_n^2 is n).

EXAMPLE 3 — T -DISTRIBUTION CI SETUP

A sample of $n = 16$ from a normal population gives $\bar{x} = 25$, $s = 4$. Construct the basis for a 95% confidence interval for μ (using $t_{15,0.025} = 2.131$).

Solution:

$$\bar{x} \pm t_{15,0.025} \frac{s}{\sqrt{n}} = 25 \pm 2.131 \times \frac{4}{4} = 25 \pm 2.131$$

Giving CI $\approx (22.87, 27.13)$.

EXAMPLE 4 — F -DISTRIBUTION RELATION

If $T \sim t_{10}$, what is the distribution of T^2 ?

Solution: By the relation $T^2 \sim F_{1,n}$ when $T \sim t_n$: here $T^2 \sim F_{1,10}$.

7. Common Pitfalls (Exam Traps)

WATCH OUT

- Use the t -distribution when σ is **unknown** and estimated by s ; use Z (normal) when σ is **known** — mixing these up is a very common error.
- Degrees of freedom for the sample variance S^2 is $n - 1$, not n — always subtract 1 for each parameter estimated.
- $\bar{X}$ and S^2 being independent is **special to the normal distribution** — do not assume this for other population distributions.
- CLT gives an **approximation** for large n ; for small samples from non-normal populations, the normal approximation for $\bar{X}$ can be poor.
- Don't confuse χ_n^2 (used for variance inference) with the χ^2 goodness-of-fit test statistic — related but applied in different contexts.

8. Quick Revision Summary

Concept	Key Fact
CLT	$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0, 1)$
χ_n^2	Sum of n squared independent $N(0, 1)$'s; mean n , var $2n$
t_n	$Z/\sqrt{W/n}$; heavier tails than normal; $\rightarrow N(0, 1)$
$F_{m,n}$	Ratio of two independent scaled χ^2 's
T^2 relation	$T \sim t_n \implies T^2 \sim F_{1,n}$
Sample variance df	$n - 1$ (one lost estimating μ)
Normal population special property	$\bar{X} \perp S^2$ (independent)

[UGC NET](#)[Statistics](#)[Sampling Distributions](#)[CLT](#)