Bessel’s Correction – Complete Conceptual Notes

These notes focus ONLY on Bessel’s correction and explain it from first principles, covering intuition, mathematics, degrees of freedom, and common doubts. They are suitable for exams, interviews, and statistical foundations.

1. What is Bessel’s Correction?

Bessel’s correction refers to dividing by $(n - 1)$ instead of n when calculating sample variance or standard deviation.

Population variance:

$\sigma^2 = (1/N)\ \Sigma\ (x_i - \mu)^2$

Sample variance:

$s^2 = (1/(n - 1))\ \Sigma\ (x_i - x\blacksquare)^2$

2. Why Bessel’s Correction is Needed

In practice, the population mean μ is unknown and is replaced by the sample mean $x\blacksquare$.

The sample mean is estimated from the same data used to compute variance.

This causes the squared deviations to be systematically smaller, leading to underestimation if divided by n.

3. Key Constraint Introduced by Sample Mean

For the sample mean $x\blacksquare$, the deviations always satisfy:

$(x_1 - x\blacksquare) + (x_2 - x\blacksquare) + \ldots + (x_n - x\blacksquare) = 0$

This equation is always true and introduces one strict constraint on the deviations.

4. Degrees of Freedom Explanation

Degrees of freedom represent the number of independent values that can vary freely.

For n observations:

• Deviations from population mean $\mu \rightarrow n$ independent deviations

• Deviations from sample mean $x\blacksquare \rightarrow n - 1$ independent deviations

One degree of freedom is lost because the sample mean is estimated from the data.

5. Why Divide by (n – 1) and Not n

Using $x\blacksquare$ instead of μ removes one independent direction of variation.

As a result, the expected sum of squared deviations is reduced by exactly one unit of variance.

Mathematically:

$E[\ \Sigma\ (x_i - x\blacksquare)^2\ ] = (n - 1)\ \sigma^2$

Dividing by n gives an expected value of $(n - 1)/n \times \sigma^2$, which underestimates variance.

Dividing by $(n - 1)$ removes this bias.

6. Why Not Divide by (n – 2)?

Dividing by (n – 2) would imply that two parameters were estimated.

In variance estimation, only ONE parameter (the mean) is estimated.

Therefore, only one degree of freedom is lost.

7. Population Mean vs Sample Mean (Critical Difference)

For population mean μ:

$\Sigma (x_i - \mu) \neq 0$ in general (zero only in expectation).

For sample mean x■:

Σ (x_i – x■) = 0 exactly for every dataset.

This exact constraint is the root cause of variance underestimation.

8. When Would (n – 2) Be Used?

When TWO parameters are estimated from data.

Example: Linear regression estimates slope and intercept.

Residual variance in regression is divided by (n – 2).

9. Key Takeaways

- Bessel’s correction applies only to variance and standard deviation.
- It corrects bias caused by estimating the mean from the data.
- One estimated parameter $\rightarrow$ subtract one degree of freedom.
- Sample mean introduces an exact constraint on deviations.
- Dividing by (n – 1) gives an unbiased estimator of population variance.

End of Notes

These notes provide a complete and focused understanding of Bessel’s correction without mixing other probability topics.