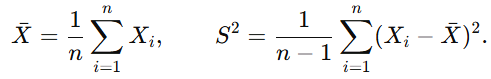
**Homework 10 — Sampling Mean/Variance and Lebesgue–Stieltjes Integration**

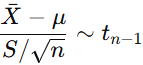
**1) Sampling mean and variance — concepts and main distributional features**

**Setup.** Let X1 ,…,Xn​ be i.i.d. with mean μ and variance σ2.

Key properties (hold under i.i.d., finite variance):

* **Unbiasedness:**
* **Precision:**
* **Consistency:**
* **Asymptotic normality:**

**Normal-data refinements (when ):**

* 
* 
* **Independence:**
* **t–statistic:**
* (Optional detail)

## 2) Lebesgue–Stieltjes (L–S) integration: idea and links to probability/measure theory

Let F:R→R be nondecreasing, right-continuous, of bounded variation.  
For measurable *f*, the **Lebesgue–Stieltjes integral** is

∫abf(x) dF(x).\int\_a^b f(x)\,dF(x).∫ab​f(x)dF(x).

**Measure-theoretic view.** FFF induces a measure μF\mu\_FμF​ on Borel sets via  
μF([a,b))=F(b)−F(a)\mu\_F([a,b))=F(b)-F(a)μF​([a,b))=F(b)−F(a). Then ∫f dF\int f\,dF∫fdF is the Lebesgue integral ∫f dμF\int f\,d\mu\_F∫fdμF​.

**Probability connection.** If FFF is the CDF of XXX, then

E[f(X)]=∫Rf(x) dF(x).\mathbb E[f(X)]=\int\_{\mathbb R} f(x)\,dF(x).E[f(X)]=∫R​f(x)dF(x).

This **unifies** discrete, continuous, and mixed laws. If FFF has density ppp on a set  
and atoms at aia\_iai​ with masses ΔF(ai)=P(X=ai)\Delta F(a\_i)=\mathbb P(X=a\_i)ΔF(ai​)=P(X=ai​), then

∫f dF=∫f(x) p(x) dx  +  ∑if(ai) ΔF(ai).\int f\,dF=\int f(x)\,p(x)\,dx\;+\;\sum\_i f(a\_i)\,\Delta F(a\_i).∫fdF=∫f(x)p(x)dx+i∑​f(ai​)ΔF(ai​).

## 3) Practice — numerical comparison: Lebesgue vs. (density-only) Riemann

**Goal.** Compute a mean/variance numerically as an L–S integral and contrast it with a “Riemann (density-only)” computation that ignores atoms.

**Test distribution (mixed):**  
With probability 0.70.70.7: U(0,1)U(0,1)U(0,1); with probability 0.30.30.3: a point mass at 222.

**Exact benchmarks (via L–S):**

E[X]=0.7⋅12+0.3⋅2=0.95,E[X2]=0.7⋅13+0.3⋅4=1.433333…\mathbb E[X]=0.7\cdot \tfrac12 + 0.3\cdot 2=0.95,\qquad \mathbb E[X^2]=0.7\cdot \tfrac13 + 0.3\cdot 4 = 1.433333\ldotsE[X]=0.7⋅21​+0.3⋅2=0.95,E[X2]=0.7⋅31​+0.3⋅4=1.433333… Var(X)=E[X2]−E[X]2=1.433333…−0.9025=0.530833…\mathrm{Var}(X)=\mathbb E[X^2]-\mathbb E[X]^2 =1.433333\ldots-0.9025=0.530833\ldotsVar(X)=E[X2]−E[X]2=1.433333…−0.9025=0.530833…

**What the density-only Riemann approach would yield (incorrectly, by ignoring the atom):**

Edens[X]=∫01x⋅0.7 dx=0.35,Vardens(X)=0.7⋅13−0.352=0.110833…\mathbb E\_{\text{dens}}[X]=\int\_0^1 x\cdot 0.7\,dx=0.35,\qquad \mathrm{Var}\_{\text{dens}}(X)=0.7\cdot\tfrac13-0.35^2=0.110833\ldotsEdens​[X]=∫01​x⋅0.7dx=0.35,Vardens​(X)=0.7⋅31​−0.352=0.110833…

**Numerical plan for the code (to be attached):**

1. Use trapezoidal/Simpson rule to approximate ∫01x (0.7) dx\int\_0^1 x\, (0.7)\,dx∫01​x(0.7)dx and ∫01x2 (0.7) dx\int\_0^1 x^2\, (0.7)\,dx∫01​x2(0.7)dx.
2. **L–S estimate:** add the atomic contributions 0.3⋅f(2)0.3\cdot f(2)0.3⋅f(2) with f(x)=xf(x)=xf(x)=x and f(x)=x2f(x)=x^2f(x)=x2.
3. **Riemann (density-only) estimate:** omit the atom to show the undercount.
4. Optionally, simulate large samples from the mixture to confirm the L–S values.

**What this demonstrates.** The L–S integral naturally captures discrete mass (∑f(ai)ΔF(ai)\sum f(a\_i)\Delta F(a\_i)∑f(ai​)ΔF(ai​)), while a density-only Riemann calculation misses it—hence the large discrepancy above.