Central limit theorem / Sampling

CLT

Remember we can standardize any Gaussian using.

$$Z = \frac{x - \mu_x}{\sigma_x} \quad \text{or} \quad Z = \frac{W_n - \mu_w}{\sigma_w}$$

With $W_n = X_1 + X_2 + \ldots + X_n = \sum_{i=1}^{n} X_i$

$$E[W_n] = \mu_w = n\mu_x$$

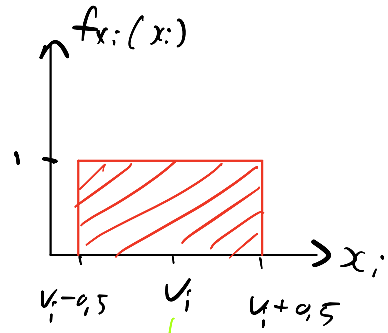
$$Var[W_n] = \sigma_w^2 = n\,Var[X] \quad \Rightarrow \quad \sigma_w = \sqrt{n\,\sigma_x^2}$$

$$= n\,\sigma_x^2$$

$$F_{W_n}(w) = P(W_n \leq w)$$

Example 9.8

- 1-bit $\Rightarrow$ Value can be 1 or 0 (X_i) — $V_i = 1/0$
- Uniformly distributed around V (± 0.5 mV)
  - $\mu_{X_i} = \dfrac{b+a}{2} = \dfrac{(V_i + 0.5) + (V_i - 0.5)}{2} = V_i$
  - $\sigma_{X_i}^2 = \dfrac{(b-a)^2}{12} = \dfrac{1}{12}.$



either 0 or 1 V

- Output is average of 8 measurements
  - $W_8 = \sum_{i=1}^{8} X_i$ — Total $\qquad U = \dfrac{W_8}{8}$ — Average.
- Question: Error $= |U - V_i|$

U: estimate (RV)

V_i: true value (constant) 0 or 1

$P(\text{Error} > 0.1) = ?$

$P(|U - V_i| > 0.1) = ?$

Solve using Central Limit Theorem approximation since iid.
(i.e. use Gaussian, we need μ and σ)

$\mu \approx 0 \qquad \sigma \approx \sqrt{\frac{1}{96}}$

$E[U - V_i] = E[U] - V_i$

$= V_i - V_i = 0$

$E[U] = E\left[\frac{W_8}{8}\right] = \dfrac{E[W_8]}{8} = \dfrac{\cancel{8}\,E[X_i]}{\cancel{8}} = V_i$

$Var[U - V_i] = Var[U] = Var\left[\frac{W_8}{8}\right]$

$= \left(\frac{1}{8}\right)^2 Var[W_8]$

$$= \left(\frac{1}{8}\right)^2 8 \cdot Var[X_i]$$

$$= \frac{1}{8} \cdot \frac{1}{12} = \frac{1}{96}$$

$$P(|u - v_i| > 0{,}1) = 1 - P(|u - v_i| \leq 0{,}1)$$

$$= 1 - P(-0{,}1 \leq u - v_i \leq 0{,}1)$$

$$1 - \Phi\left(\frac{0{,}1}{\sqrt{\frac{1}{96}}}\right)$$

$$= 1 - P(-0{,}1 \leq u \leq 0{,}1) \qquad v_i = 0$$

$$= 1 - \left[\Phi\left(\frac{0{,}1 - 0}{\sqrt{\frac{1}{96}}}\right) - \Phi\left(\frac{-0{,}1 - 0}{\sqrt{\frac{1}{96}}}\right)\right]$$

$$= 2\left[1 - \Phi\left(\underbrace{\frac{0{,}1}{\sqrt{\frac{1}{96}}}}_{0{,}98}\right)\right]$$

$$= 0{,}3272$$

Sample mean

$$E[M_n(X)] = E[X] = \mu_X$$

$$Var[M_n(X)] = \frac{Var[X]}{n}$$

Proof: (for iid)

Remember:

$$Var[X_1 + X_2 + \dots + X_n] = Var[X_1] + Var[X_2] + \dots + Var[X_n]$$
$$= n\, Var[X]$$

$$Var[aX] = a^2\, Var[X]$$

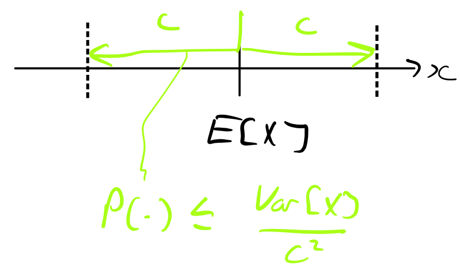
$$Var[M_n(X)] = Var\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right]$$
$$= \frac{1}{n^2} \cdot n\, Var[X] = \frac{Var[X]}{n}$$

Markov inequality

$$P(X \geq c) \leq \frac{E[X]}{c}$$

Chebyshev Inequality

$$P(|X - \mu_X| \geq c) \leq \frac{Var[X]}{c^2}$$



Proof:

- Set $c_{old} = c^2$ in Markov ineq.

$$P(X \geq c^2) \leq \frac{E[X]}{c^2}$$

- Set $X = (Y - \mu_Y)^2$

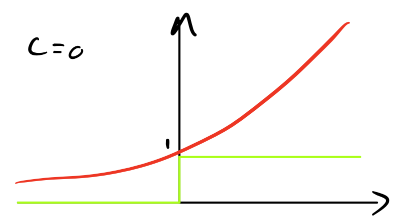
$$P((Y-\mu_Y)^2 \geq c^2) \leq \frac{E[(Y-\mu_Y)^2]}{c^2} = \frac{Var[Y]}{c^2}$$

$$\Rightarrow P(|Y - \mu_Y| \geq c) \leq \frac{Var[Y]}{c^2}$$

Chernoff Bound

$$P(X \geq c) = \int_c^\infty f_X(x)\,dx = \int_{-\infty}^{\infty} \underbrace{u(x-c)}_{\text{Unit step at } x=c} f_X(x)\,dx$$

Now note for $s \geq 0$ $\quad u(x-c) \leq e^{s(x-c)}$



$$\therefore P(X \geq c) \leq \int_{-\infty}^{\infty} e^{s(x-c)} f_X(x)\,dx$$

$$= e^{-sc} \int_{-\infty}^{\infty} e^{sx} f_X(x)\,dx$$

$$= e^{-sc} \phi_X(s)$$

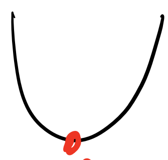
for the tightest bound minimize s (but $s \geq 0$)

Example 10.4

$$\min_{s \geq 0} (s^2 - 11s)$$

$$2s - 11 = 0$$

$$s = \frac{11}{2}$$



where $\frac{d}{ds} = 0$