r.v.

$$X : \Omega \longrightarrow \mathbb{R}$$

probability mass function (density) p.m.f.

discrete $p_X(x) = \Pr(X = x)$

$x \in \mathbb{R}$

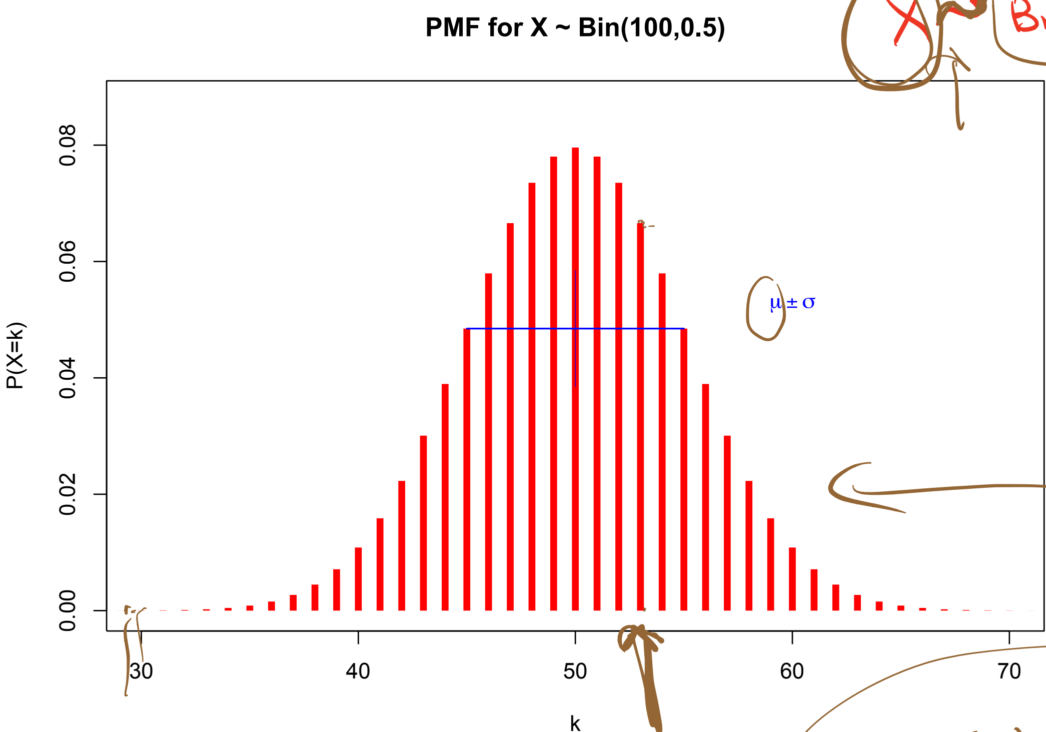
$$X = \begin{cases} 1 & \frac{3}{4} \\ 0 & \frac{1}{4} \end{cases}$$

$p_X(1) = \frac{3}{4}$

$p_X(2) = 0$

binomial random variable

$X \sim \mathrm{Bin}(n, p)$



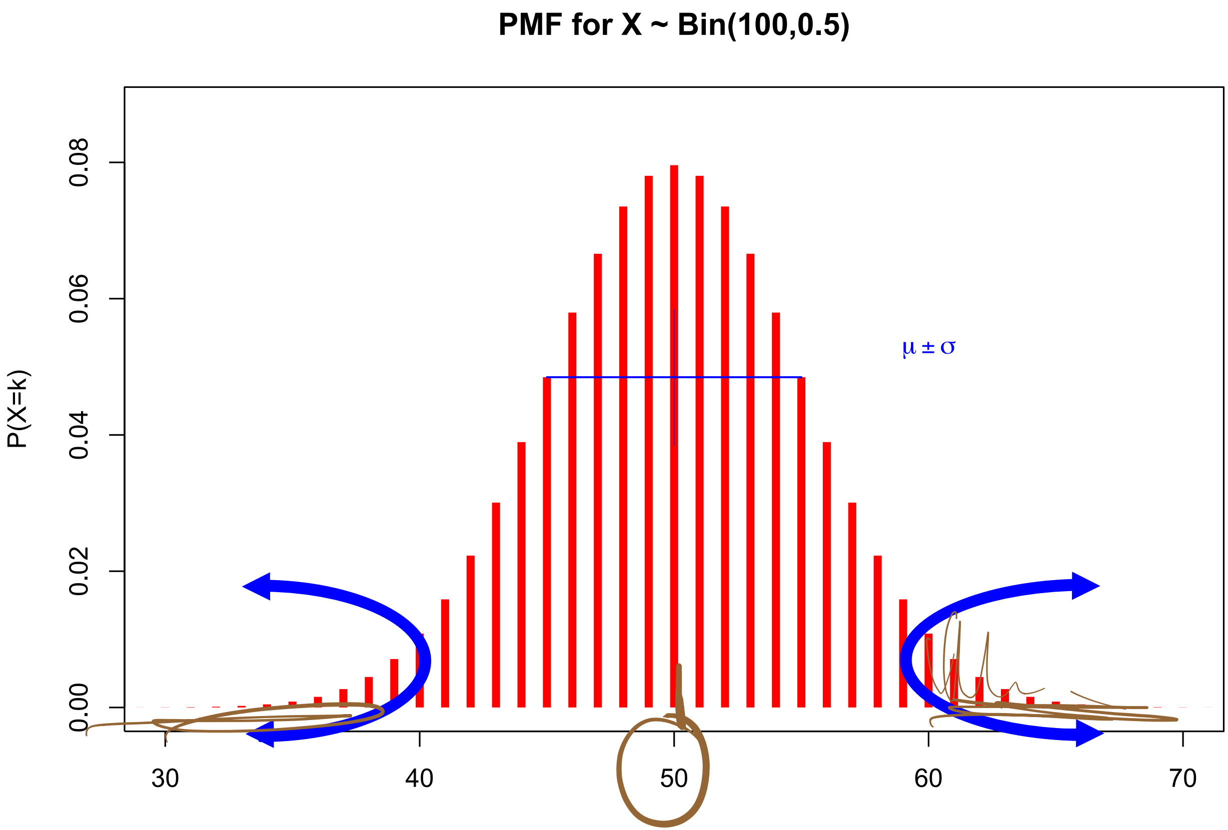
X counts the # of H's in n indep coin tosses each with prob p of coming up H's

p.m.f for Bin(100, 0.5)

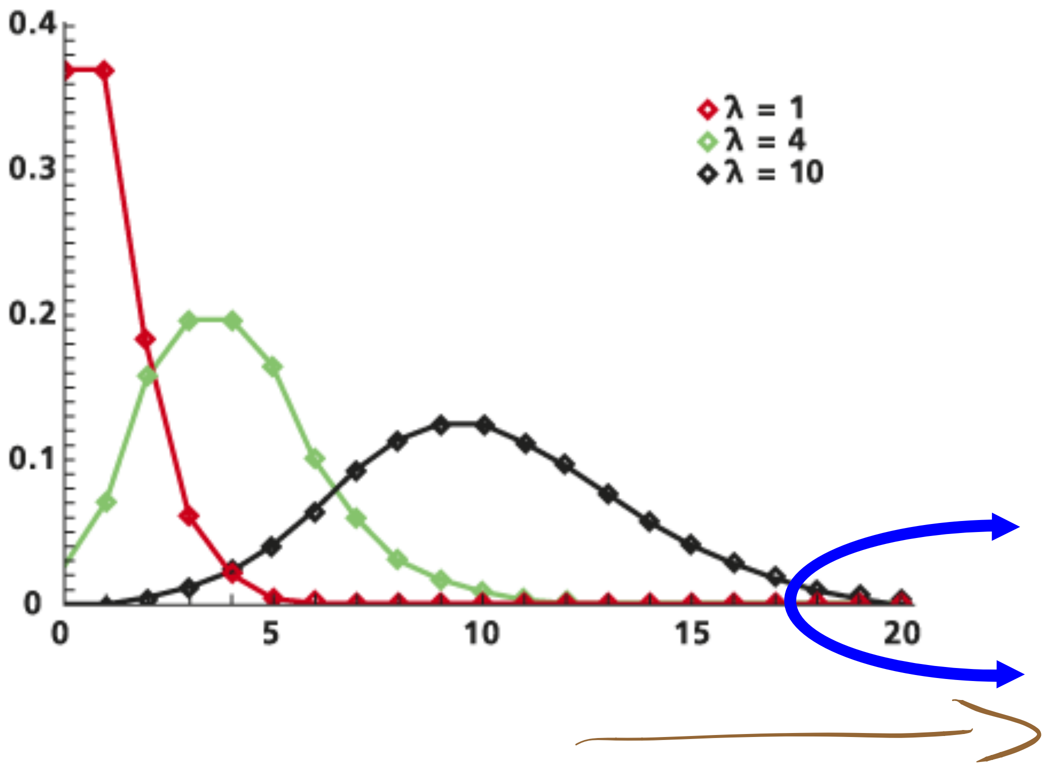
$k = 53$

$$E(X) = \sum_{k=0}^{100} k \, p_X(k)$$

For a random variable X, the *tails* of X are the parts of the PMF that are "far" from its mean.



For a random variable X, the *tails* of X are the parts of the PMF that are "far" from its mean.



Often, we want to bound the probability that a random variable X is "extreme."

Such a bound is called a "tail bound".

Suppose we know that X is always non-negative.

Theorem: If X is a non-negative random variable, then for every α > 0, we have

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

Equivalently

$$P(X \geq cE[X]) \leq 1/c$$

100 $\leq \frac{1}{100}$

Theorem: If X is a non-negative random variable, then for every $\alpha > 0$, we have

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

Proof:

$$E[X] = \Sigma_x \, x \, p(x)$$

$$= \Sigma_{x<\alpha} \, x \, p(x) + \Sigma_{x\geq\alpha} \, x \, p(x)$$

$$\geq \quad \odot \quad + \Sigma_{x\geq\alpha} \alpha p(x)$$ $(x \geq 0;\ \alpha \leq x)$

$$= \alpha P(X \geq \alpha)$$

If we know *more* about a random variable, we can often use that to get *better* tail bounds.

Suppose we *also* know the variance.

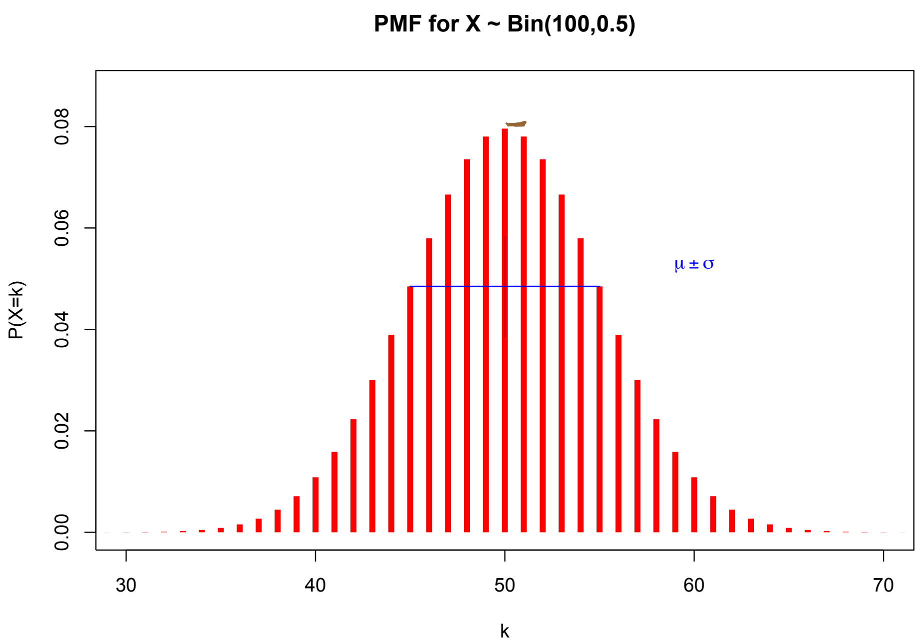
$E((Y-\mu)) = 0$ $\qquad$ $E(Y) - E(\mu) = 0$

$$\mathrm{Var}[Y] = E[(Y - E(Y))^2] = \sigma^2$$

Standard deviation = square root of variance $= \sigma$

$$\mathrm{Var}[Y] = E[(Y - \underbrace{E(Y)}_{\mu})^2] = E[(Y-\mu)^2]$$

25



$\sigma = 5$

If we know *more* about a random variable, we can often use that to get *better* tail bounds.

Suppose we *also* know the variance.

Theorem: If Y is an arbitrary random variable with E[Y] = μ, then, for any α > 0,

$$P(|Y - \mu| \geq \alpha) \leq \frac{\mathrm{Var}[Y]}{\alpha^2}$$

$$\boxed{\alpha = c\sigma} \quad \leq \frac{\sigma^2}{(c\sigma)^2} = \frac{1}{c^2}$$

$$\Pr(|Y - \mu| \geq c\sigma) \leq \frac{1}{c^2}$$

Theorem: If Y is an arbitrary random variable with μ = E[Y], then, for any α > 0,

$$P(|Y - \mu| \geq \alpha) \leq \frac{\mathrm{Var}[Y]}{\alpha^2}$$

Proof: $\text{Let } X = (Y - \mu)^2$

X is non-negative, so we can apply Markov's inequality:

$$P(|Y - \mu| \geq \alpha) = P(X \geq \alpha^2)$$

$$\leq \frac{E[X]}{\alpha^2} = \frac{\mathrm{Var}[Y]}{\alpha^2}$$

$\mathrm{Var}[Y] = E[(Y - E(Y))^2]$

Important Facts about Variance

Linearity of expectation always holds, i.e.,

$$E(a_1 X_1 + a_2 X_2 + \ldots + a_n X_n) = a_1 E(X_1) + a_2 E(X_2) + \ldots + a_n E(X_n)$$

Linearity of variance holds **only** if the random variables are independent.

$$\mathrm{Var}[X_1 + \ldots + X_n] = \mathrm{Var}[X_1] + \ldots + \mathrm{Var}[X_n]$$

Also, if a is a constant, then

$$\mathrm{Var}[a \cdot X] = a^2 \cdot \mathrm{Var}[X]$$

def X, Y indep

$\equiv \Pr(X=a, Y=b) = \Pr(X=a)\Pr(Y=b)$

Performance/estimation using ~~repetition~~ independent

Example: $X = \begin{cases} 1 & p \\ 0 & 1-p \end{cases}$

$X_1 = 1$ $\quad \frac{p}{1}$

$X_2 = 1$ $\quad 1$

$X_3 = 0$ $\quad \frac{2}{3}$

n trials, estimate $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{n\to\infty} p$

$\mathrm{Var}(X_i) = \sigma^2$

$$\boxed{\bar{X}_n = \frac{1}{n}(X_1 + X_2 + \dots + X_n)}$$

$$E(\bar{X}_n) = \frac{1}{n}\sum_{i=1}^{n} E(X_i)$$

$$= \frac{1}{n}\cdot n\mu = \mu$$

$$\mathrm{Var}(\bar{X}_n) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$$

$$= \frac{1}{n^2}\underbrace{\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right)}_{\text{indep}}$$

$$\underset{\text{lin of variance for indep}}{=} \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

Law of large Numbers

$X_1, X_2, \dots, X_n$ all from same distn, i.i.d. with mean μ

$$\boxed{\lim_{n\to\infty} \underbrace{\frac{1}{n}\sum_{i=1}^{n} X_i}_{\bar{X}_n} \rightarrow E(X_i) = \mu}$$

$$\Pr(|\bar{X}_n - \mu| \geq c) \leq \frac{\mathrm{Var}(\bar{X}_n)}{c^2}$$

$$= \frac{\sigma^2}{n c^2} \xrightarrow{n\to\infty} 0$$