Variance.

Definition. Let X have pmf $p(x)$ and expected value μ. Then the variance of X, denoted by $V(X)$ or σ_X^2, is

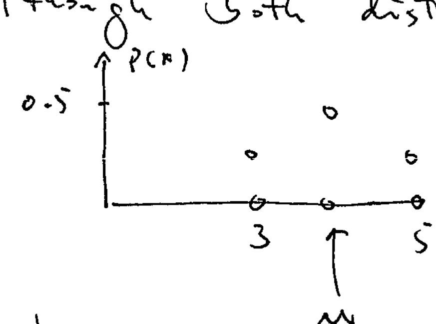
$$\sigma_X^2 = V(X) = E\left[(X-\mu)^2\right] = \sum_x (x-\mu)^2 p(x)$$

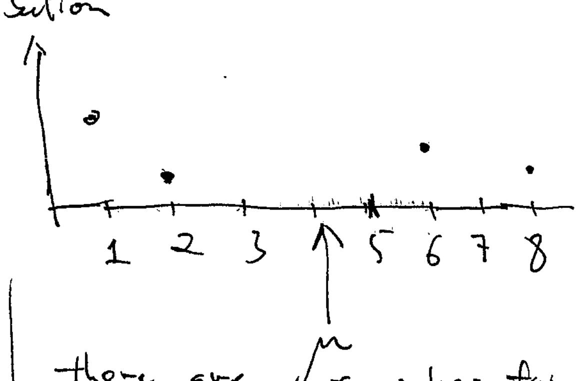
The standard deviation (SD) of X is

$$\sigma_X = \sqrt{\sigma_X^2}$$

Remark. - $h(X) = (X-\mu)^2$ is the squared deviation of X from its mean, and σ^2 is the expected value of the squared deviatiation.

- Although both distribution





most of the probability distribution is close to μ, then σ^2 will be relatively small.

there are x values far from μ that have large $p(x)$, then σ^2 will be large.

Proposition (Shortcut formula for σ^2):

$$\mathbb{V}(\underline{X}) = E(\underline{X}^2) - [E(X)]^2 = \sum_x x^2 p(x) - \mu.$$

Proof.

$$\mathbb{V}(\underline{X}) = \sum_x (x-\mu)^2 p(x) = \sum_x x^2 p(x) - 2\mu\underbrace{\left(\sum_x x p(x)\right)}_{=\mu} + \mu^2 \underbrace{\left(\sum_x p(x)\right)}_{=1}$$

$$= E(\underline{X}^2) - 2\mu^2 + \mu^2$$

$$= E(\underline{X}^2) - \mu^2$$

Variance of function.

Proposition:

$$\mathbb{V}[h(\underline{X})] = \sigma^2_{h(\underline{X})} = \sum_x \left(h(x) - E[h(\underline{X})]^2\right) p(x).$$

Proof: By definition,

$$\sigma^2_{h(\underline{X})} = E\left[\left(h(\underline{X}) - \mu_{h(\underline{X})}\right)^2\right]$$

$$= \sum_x \left(h(x) - E[h(\underline{X})]\right)^2 p(x)$$

Proposition:

$$\mathbb{V}(a\underline{X}+b) = a^2\, \mathbb{V}(\underline{X})$$

Proof:

$$\mathbb{V}(a\underline{X}+b) = \sum_x \left(ax + b - E(a\underline{X}+b)\right)^2 p(x)$$

$$= \sum_x \left(ax + b - aE(\underline{X}) - b\right)^2 p(x)$$

$$= a^2 \sum_x (x - E(\underline{X}))^2 p(x)$$

$$= \sigma^2\, \mathbb{V}(\underline{X}).$$

Moments

Definition. – Expected values of powers of X are called moments about 0.

– Expected values of powers of $X - \mu$ are called moments about the mean.

Example. – The first moment about 0 of X is the mean:

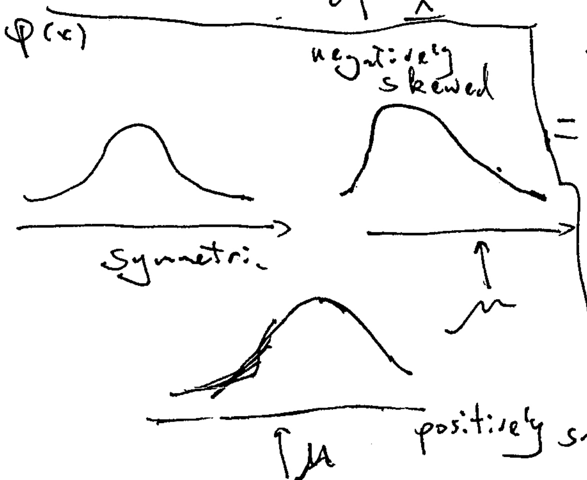
$$\mu = E(X) = \sum_x x\varphi(x)$$

– The second moment about the mean is the variance:

$$\sigma^2 = E\left[(X-\mu)^2\right] = \sum_x (x-\mu)^2\varphi(x)$$

– The third moment about the mean divided by σ^3 is the skewness

$$\text{skewness of } X = \frac{E\left[(X-\mu)^3\right]}{\sigma^3} = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right]$$



$$= \begin{cases} 0 & \text{if the distribution } \varphi(x) \text{ is symmetric} \\ <0 & \text{if the distribution } \varphi(x) \text{ is skewed to the left} \Rightarrow \text{negatively skewed} \\ >0 & \text{if the distribution } \varphi(x) \text{ is skewed to the right} \Rightarrow \text{positively skewed} \end{cases}$$

Moment Generating Function

Definition: The moment generating function (mgf) of $\underline{X}$ is defined to be

$$M_{\underline{X}}(t) = E\left(e^{t\underline{X}}\right) = \sum_{x} e^{tx} p(x)$$

Remark – We assume that $M_{\underline{X}}(t)$ is defined for an interval including 0 in its interior

– $M_{\underline{X}}(0) = \sum_{x} p(x) = 1$.

<u>Theorem</u>: For any positive integer k, let

$$M_{\underline{X}}^{(k)}(t) = \frac{d^k}{dt^k}\left[M_{\underline{X}}(t)\right]$$

denote the kth derivative of $M_{\underline{X}}(t)$. Then

The k^{th} moment about 0 $= E\left(\underline{X}^k\right) = M_{\underline{X}}^{(k)}(0)$.

Proof: By induction, show $M_{\underline{X}}^{(k)}(t) = \sum_{x} x^k e^{tx} p(x)$

– $k=1$. Differentiate, $\forall t$ inside the interval of convergence,

$$M_{\underline{X}}^{(1)}(t) = \frac{d}{dt} M_{\underline{X}}(t) = \frac{d}{dt}\left[\sum_{x} e^{tx} p(x)\right] = \sum_{x} \frac{d}{dt}\left[e^{tx}\right] p(x)$$

$$= \sum_{x} x e^{tx} p(x)$$

Assume that for $k-1$ we have $M_{\underline{X}}^{(k-1)}(t) = \sum_{x} x^{k-1} e^{tx} p(x)$

Then

$$M_{\underline{X}}^{(k)}(t) = \frac{d}{dt}\left[M_{\underline{X}}^{(k-1)}(t)\right] = \frac{d}{dt} \sum_{x} x^{k-1} e^{tx} p(x) = \sum_{x} x^{k-1} \frac{d}{dt}\left[e^{tx}\right] p(x)$$

$$= \sum_{x} x^k e^{tx} p(x)$$

* Set $t=0$

$$M_X^{(k)}(0) = \sum_x x^k p(x) = E[X^k].$$

Example: Let X have pmf

x	0	1	2
$p(x)$	0.7	0.2	0.1

Then $M_X(t) = \sum_x e^{tx} p(x)$

$$= p(0) + e^t p(1) + p(2)e^{2t}$$

$$= 0.7 + 0.2e^t + 0.1e^{2t}$$

First,

$$M_X^{(1)}(t) = 0.2e^t + 0.1\,(2)\,e^{2t}$$

$$M_X^{(2)}(t) = 0.2e^t + 0.1\,(2)(2)\,e^{2t}$$

Then setting $t=0$ gives

$$\mu = E(X) = M_X^{(1)}(0) = 0.2 + 0.1\,(2) = 0.4$$

$$E(X^2) = M_X^{(2)}(0) = 0.2 + 0.1(2)(2) = 0.6$$

Then using shortcut formula for σ^2 gives

$$\sigma^2 = E(X^2) - [E(X)]^2 = 0.6 - (0.4)^2 = 0.44$$

Example: * Let X be a geometric rv with pmf

$$p(j) = p(1-p)^{j-1}, \quad j = 1, 2, 3, \ldots$$

Then the mgf is

$$M_X(t) = \sum_x e^{tx} p(x) = \sum_{j=1}^{\infty} e^{tj} p(1-p)^{j-1}$$

$$= pe^t \sum_{j=1}^{\infty} \left[e^t(1-p)\right]^{j-1}$$

$$= \frac{pe^t}{1-(1-p)e^t}$$

which is defined for t s.t

$$|e^t(1-p)| < 1$$

implying

$$t < -\ln(1-p).$$

* Checking that $M_X(\underset{0}{\overset{\shortparallel}{t}}) \overset{?}{=} 1$

$$= \frac{p}{1-(1-p)} = 1 \ (\checkmark \text{ ok}).$$

* Differentiate

$$M_X^{(1)}(t) = \frac{pe^t}{[1-(1-p)e^t]^2}$$

$$M_X^{(2)}(t) = \frac{pe^t[1+(1-p)e^t]}{[1-(1-p)e^t]^3}$$

* Setting $t=0$ gives

$$\mu = E(X) = M_X^{(1)}(0) = \frac{1}{p}, \quad E(X^2) = M_X^{(2)}(0) = \frac{2-p}{p^2}$$

thus

$$\sigma^2 = E(X^2) - [E(X)]^2 = \frac{1-p}{p^2}$$

Proposition. (shortcut formula).

Let define $R_{\bar{X}}(t) = \ln\left[M_{\bar{X}}(t)\right]$. Then

$$\mu = E(\bar{X}) = R_{\bar{X}}^{(1)}(0)$$

$$\sigma^2 = V(\bar{X}) = R_{\bar{X}}^{(2)}(0)$$

Proof. First,

$$R_{\bar{X}}^{(1)}(t) = \frac{d}{dt}\left[\ln\left(M_{\bar{X}}(t)\right)\right] = \frac{1}{M_{\bar{X}}(t)} M_{\bar{X}}^{(1)}(t)$$

$$R_{\bar{X}}^{(2)}(t) = \frac{d}{dt}\left[R_{\bar{X}}^{(1)}(t)\right] = \frac{d}{dt}\left[\frac{1}{M_{\bar{X}}(t)} M_{\bar{X}}^{(1)}(t)\right]$$

$$= \frac{1}{M_{\bar{X}}(t)} M_{\bar{X}}^{(2)}(t) - \frac{1}{\left[M_{\bar{X}}(t)\right]^2}\left[M_X^{(1)}(t)\right]^2$$

Setting $t=0$ gives

$$R_{\bar{X}}^{(1)}(t=0) = \frac{1}{\underbrace{M_{\bar{X}}(0)}_{=1}} \underbrace{M_{\bar{X}}^{(1)}(0)}_{=\mu} = \mu$$

$$R_{\bar{X}}^{(2)}(t=0) = \frac{1}{\underbrace{M_{\bar{X}}(0)}_{=1}} \underbrace{M_{\bar{X}}^{(2)}(t=0)}_{=E[\bar{X}^2]} - \frac{1}{\underbrace{\left[M_X(0)\right]^2}_{=1}}\Big[\underbrace{M_X^{(1)}(0)}_{=\mu}\Big]^2$$

$$= E\left[\bar{X}^2\right] - \left[E(X)\right]^2$$

$$= \sigma^2.$$

Example: Let $\underline{X}$ be geometric with

$$\varphi(j) = p(1-p)^{j-1}, \quad j = 1, 2, \ldots$$

Then

$$M_{\underline{X}}(t) = \frac{p e^t}{1-(1-p)e^t}$$

gives

$$R_{\underline{X}}(t) = \ln\left[M_{\underline{X}}(t)\right]$$

$$= (\ln p + t) - \ln\left[1-(1-p)e^t\right]$$

Then

$$R_{\underline{X}}^{(1)}(t) = \frac{1}{1-(1-p)e^t}$$

$$R_{\underline{X}}^{(2)}(t) = \frac{(1-p)e^t}{\left[1-(1-p)e^t\right]^2}$$

Setting $t=0$ gives

$$\begin{cases} \mu = R_{\underline{X}}^{(1)}(0) = \dfrac{1}{p} \\ \sigma^2 = R_{\underline{X}}^{(2)}(0) = \dfrac{1-p}{p^2} \end{cases}$$

proposition:

$$M_{[a\underline{X}+b]}(t) = e^{bt} M_{\underline{X}}(at)$$

Proof.

$$M_{[a\overline{X}+b]}(t) = E\left[e^{t(a\overline{X}+b)}\right]$$

$$= E\left[e^{bt} e^{at\overline{X}}\right]$$

$$= e^{bt} E\left[e^{(at)\overline{X}}\right]$$

$$= e^{bt} M_{\underline{X}}(at).$$

Example: Let $Y = 10\overline{X} - 5$ and $\overline{X}$ have p.m.f. Then

x	0	1
$p(x)$	$\frac{20}{38}$	$\frac{18}{18}$

$y = 10x - 5$	-5	5
$p(y)$	$\frac{20}{38}$	$\frac{18}{38}$

$$M_{\overline{Y}}(t) = \sum_y e^{ty} p(y) = e^{-5t}\left(\frac{20}{38}\right) + e^{5t}\left(\frac{18}{38}\right)$$

$$\shortparallel$$

$$M_{[10\overline{X}-5]} = e^{-5t} M_x(10t) = e^{-5t}\left[\sum_x e^{10tx} p(x)\right]$$

$$= e^{-5t}\left[\frac{20}{38} + e^{10t}\left(\frac{18}{38}\right)\right]$$

$$= e^{-5t}\left(\frac{20}{38}\right) + e^{5t}\left(\frac{18}{38}\right).$$