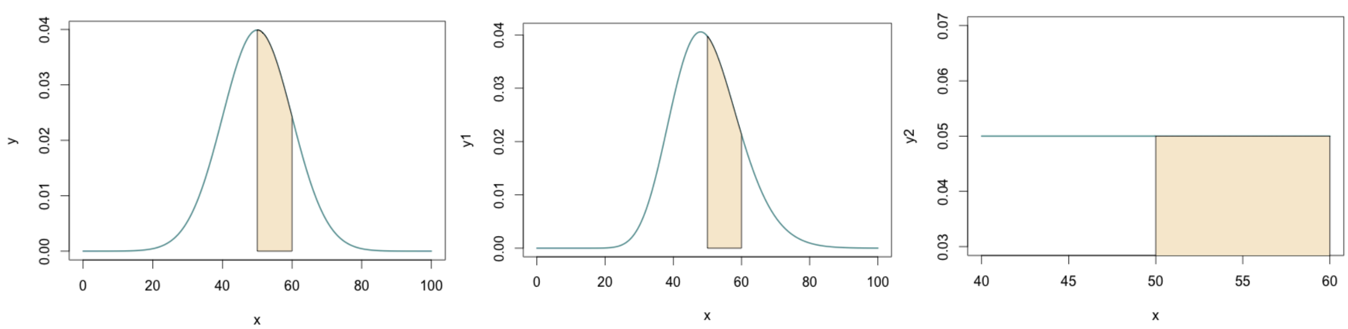
A desigualdade de Tchebycheff:

A desigualdade proposta pelo pesquisador russo Pafnuty Lvovich Tchebycheff fornece meios para compreender como a variância mede a variabilidade em relação ao valor esperado. Vejamos como isso ocorre

Se conhecermos a distribuição de probabilidade podemos calcular $E(X)$ e $Var(X)$. No entanto, se conhecermos $E(X)$ e $Var(X)$, não é possível reconstruir a distribuição de probabilidade. Dessa forma, sabendo apenas a variância e esperança não podemos calcular $P(|X - E(X)| \leq c)$, onde c é um valor pequeno qualquer.

A figura abaixo representa três possíveis distribuições, sendo que todas possuem a mesma esperança (50) e desvio padrão (10). Observe que a $P(|X - E(X)| \leq c)$ é diferente para cada uma delas. Assim, sem saber qual é a distribuição não se pode calcular a probabilidade acima.

Figura – Distribuição normal, chi-quadrada e uniforme com esperança igual a 50 e desvio padrão igual a 10. Área marcada define a probabilidade $P(X - E(X))$



(a) (b) (c)

Apesar da impossibilidade de calcular $P(|X - E(X)| \leq c)$ é possível estabelecer limites superiores e inferiores para a variabilidade ao redor do valor esperado.

PROGRAMAÇÃO EM R:

Distribuição Normal

```
x=seq(00,100,length=200)
 y=dnorm(x,mean=50,sd=10)
 plot(x,y,type="l",lwd=2,col="cadetblue")
 x=seq(50,60,length=200)
 y=dnorm(x,mean=50,sd=10)
 polygon(c(50,x,60),c(0,y,0),col="antiquewhite")
```

Distribuição Uniforme

```
x=seq(40,60,length=200)
 y2=dunif(x, min=40,max=60)
 plot(x,y2,type="l",lwd=2,col="cadetblue")
 x=seq(50,60,length=200)
 y2=dunif(x, min=40,max=60)
 polygon(c(50,x,60),c(0,y2,0),col="antiquewhite")
```

Distribuição Chi Quadrada

```
x=seq(0,100,length=200)
 y1=dchisq(x,df=50)
 plot(x,y1,type="l",lwd=2,col="cadetblue")
 x=seq(50,60,length=200)
 y1=dchisq(x,df=50)
 polygon(c(50,x,60),c(0,y1,0),col="antiquewhite")
```

<u>Definição:</u>

Seja X uma variável aleatória com $E(X) = \mu$ e c um número real. Se $E(X - c)^2$ for finita e ε for qualquer número positivo, tem-se:

$$P[|X - c| \geq \varepsilon] \leq \frac{1}{\varepsilon^2} E(X - c)^2$$

Formas equivalentes

(1) Complementar

$$P[|X - c| \geq \varepsilon] \geq 1 - \frac{1}{\varepsilon^2} E(X - c)^2$$

(2) Para $c = \mu$

$$P[|X - \mu| \geq \varepsilon] \leq \frac{1}{\varepsilon} E(X - \mu)^2 = P[|X - \mu| \geq \varepsilon] \leq \frac{\sigma^2}{\varepsilon}$$

(3) Para $c = \mu$ e $\varepsilon = k\sigma$

$$P[|X - \mu| \geq k\sigma] \leq \frac{1}{k^2 \cancel{\sigma^2}} \cancel{\sigma^2}$$

$$P[|X - \mu| \geq k\sigma] \leq k^{-2}$$

Demonstração:

$$\sigma^2 = Var(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx$$

$$\sigma^2 = \int_{-\infty}^{\mu-k\sigma} (x-\mu)^2 f(x)dx + \int_{\mu-k\sigma}^{\mu+k\sigma} (x-\mu)^2 f(x)dx + \int_{\mu+k\sigma}^{\infty} (x-\mu)^2 f(x)dx$$

Os três termos são negativos. Retirando o termo do meio:

$$\sigma^2 \geq \int_{-\infty}^{\mu-k\sigma} (x-\mu)^2 f(x)dx + \int_{\mu+k\sigma}^{\infty} (x-\mu)^2 f(x)dx$$

Note que para a 1ª integral

$$x \leq \mu - k\sigma = \; k\sigma \leq \mu - x = -k\sigma \geq x - \mu = (x-\mu)^2 \geq k^2\sigma^2.$$

Do mesmo modo para o 2º intervalo

$$x \;\geq\; \mu + k\sigma \;\rightarrow (x-\mu) \;\geq k\sigma \rightarrow (x\;-\mu)^2 \;\geq k^2\sigma^2$$

Assim, continua válido:

$$\sigma^2 \geq \int_{-\infty}^{\mu-k\sigma} k^2\sigma^2 f(x)dx + \int_{\mu+k\sigma}^{\infty} k^2\sigma^2 f(x)dx$$

Dividindo tudo por $k^2\sigma^2$:

$$\frac{1}{k^2} \geq \int_{-\infty}^{\mu-k\sigma} f(x)dx + \int_{\mu+k\sigma}^{\infty} f(x)dx$$

$$\frac{1}{k^2} \geq P[X \leq \mu - k\sigma] + P[X \geq \mu + k\sigma]$$

$$\frac{1}{k^2} \geq P[X - \mu \leq -k\sigma] + P[X - \mu \geq k\sigma]$$

$$\therefore P[|X-\mu| \geq k\sigma] \leq \frac{1}{k^2}$$

Note que precisa-se de muito pouco para saber do comportamento probabilístico da variável aleatória X. Caso haja informações sobre a distribuição de probabilidade da variável aleatória X possibilitará melhora na desigualdade.

EXEMPLO:

Parte 1)

Qual a probabilidade de uma variável aleatória X estar entre dois desvios padrões abaixo e dois desvio padrões acima da média:

Retomemos um passo da demonstração:

$$\frac{1}{k^2} \geq P[X \leq \mu - k\sigma] + P[X \geq \mu + k\sigma]$$

Como queremos avaliar dois desvios padrões acima e abaixo da média $k = 2$ e utilizamos o complementar:

$$1 - \frac{1}{2^2} \leq 1 - \ (P[X \leq \mu - 2\sigma] + P[X \geq \mu + 2\sigma])$$

$$0.75 \leq 1 - \ (P[X \leq \mu - 2\sigma] + P[X \geq \mu + 2\sigma])$$

Note que isso é verdade para qualquer distribuição, ou seja, a probabilidade de a variável aleatória X estar entre dois desvios padrões abaixo e dois desvio padrões acima da média é maior que 0.75.

Parte 2)

Supondo que $E(X) = 75$ e $DP(X) = 5$ qual será o limite superior e inferior para uma variável aleatória X que desvie entre dois desvios padrões abaixo e dois desvio padrões acima da média?

$$0.75 \leq 1 - \ (P[X \leq \mu - 2\sigma] + P[X \geq \mu + 2\sigma])$$

$$0.75 \leq 1 - \ (P[X \leq 75 - 10] + P[X \geq 75 + 10])$$

$$0.75 \leq 1 - \ (P[X \leq 65] + P[X \geq 85])$$

$$0.75 \leq\ P[65 \leq X \leq 85]$$

Ou seja, no mínimo 75% dos valores estão entre 65 e 85.

Parte 3)

Supondo que $E(X) = 75$ e $DP(X) = 5$ qual será a probabilidade de estar entre 83 de limite superior e 67 de limite inferior?

$$0.75 \leq 1 -\ (P[X \leq \mu - 2\sigma] + P[X \geq \mu + 2\sigma])$$

Assim:

$$\mu - k\sigma = 67 \rightarrow -k = \frac{67 - 75}{5} \rightarrow -k = \frac{67 - 75}{5} \rightarrow k = 1.6$$

$$\mu + k\sigma = 83 \rightarrow -k = \frac{83 - 75}{5} \rightarrow -k = \frac{83 - 75}{5} \rightarrow k = -1.6$$

$$1 - \frac{1}{1.6^2} \leq 1 -\ (P[X \leq \mu - 1.6\sigma] + P[X \geq \mu + 1.6\sigma])$$

ou

$$0.609 \leq P[67 \leq X \leq 83]$$

Assim, 60.9% dos valores estarão entre 67 e 83.

Retomando a desigualdade de Tchebycheff:

$P[|X - \mu| \geq k\sigma] \leq k^{-2}$

Se $Var(X)$ for pequena a maior parte da probabilidade de X estará “concentrada” próxima de $E(X)$. Vejamos mais um exemplo.

EXEMPLO:

Suponha que se quer avaliar a desigualdade para o caso onde a variável aleatória X esta entre dois desvios padrões abaixo e dois desvio padrões acima da média. Suponha que a $E(X) = 0$ e que os $DP(X) = 1, 0.5\ e\ 0.1$

$$P[|X - \mu| \geq k\sigma] \leq k^{-2}$$

$$P[|X - 0| \geq 1\sigma] \leq 1/2^2$$

$$P[|X| \geq \sigma] \leq 0.25$$

Utilizando o complementar, para os três casos tem-se:

$$P[-1 \leq X \leq 1] \geq 0.75$$

$$P[-0.5 \leq X \leq 0.5] \geq 0.75$$

$$P[-0.1 \leq X \leq 0.1] \geq 0.75$$

Percebe-se que X vai de concentrando ao redor de $E(X)$

Em razão dessa concentração segue-se o teorema:

Teorema:

Seja X uma variável aleatória com $\mathrm{V}ar(X) = 0$. Então $P[X = \mu] = 1$ ou $\mu = E(X)$

Demonstração:

$$P[|X - \mu| \geq \varepsilon] = 0 \text{ para } \forall\ \varepsilon > 0$$

$$P[|X - \mu| < \varepsilon] = 1 \text{ para } \forall\ \varepsilon > 0$$

Sendo que ε pode ser arbitrariamente pequeno.