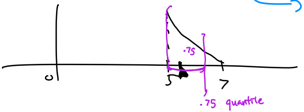
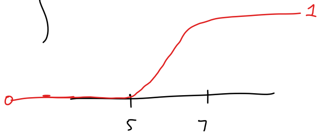
4.33

$f(y) = \frac{3}{8}(7-y)^2 \qquad \underbrace{5 \le y \le 7}$ → "Support"



$$F(y) = \begin{cases} 0 & y < 5 \\ \boxed{\frac{3}{8}\int_5^y (7-x)^2\,dx} & 5 \le y < 7 \\ 1 & y \ge 7 \end{cases}$$

continuous



$$E(Y) = \int_5^7 y \cdot \frac{3}{8}(7-y)^2\,dy = 5.5$$

$$V(X) = \int_5^7 y^2\,\frac{3}{8}(7-y)^2\,dy - 5.5^2 = .15$$

b) Find an interval in which at least 3/4 of PH measurement must lie.. Tchebyshev

$\mu \pm 2\sigma = (4.725,\ 6.275)$

Since $Y \ge 5$, $(5,\ 6.275)$

Compare this with

$[5, \phi_{.75}]$

.75 quantile

c) Would you expect to see

Ph level < 5.5 very often?

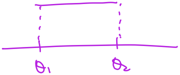
$$P(Y < 5.5) = \int_5^{5.5} \frac{3}{8}(7-y)^2 \, dy = \underline{\underline{.578}} .$$

Uniform Probability Distribution

Flat

Y has a continuous uniform probability distribution on the interval (θ_1, θ_2) iff the density of Y is

$$f(y) = \frac{1}{\theta_2 - \theta_1} \quad \theta_1 \leq y \leq \theta_2,$$



denoted as $\boxed{Y \sim U(\theta_1, \theta_2).}$

Check!!

$$\mu = E(Y) = \frac{\theta_1 + \theta_2}{2} \quad \text{and} \quad \sigma^2 = V(Y) = \frac{(\theta_2 - \theta_1)^2}{12}.$$

$$\text{MGF} = \begin{cases} \dfrac{e^{t\theta_2} - e^{t\theta_1}}{(\theta_2 - \theta_1)t} & t \neq 0 \\ 0 & t = 0 \end{cases}$$

Proof of Tchebyshev's inequality ⟶ (Markov Ineq.)

$\forall k > 0, \qquad P(|Y-\mu| \geq k\sigma) \leq \frac{1}{k^2}$

☺ $LHS = P(Y \leq \mu - k\sigma \text{ or } Y \geq \mu + k\sigma)$

$$= \int_{-\infty}^{\mu-k\sigma} 1 \cdot f(y)\,dy + \int_{\mu+k\sigma}^{\infty} f(y)\,dy$$

$y \leq \mu - k\sigma$

$\Leftrightarrow \; y-\mu \leq -k\sigma$

$\Leftrightarrow \; \frac{y-\mu}{k\sigma} \leq -1$

$\Leftrightarrow \; \frac{(y-\mu)^2}{k^2\sigma^2} \geq 1$

$$\leq \int_{-\infty}^{\mu-k\sigma} \frac{(y-\mu)^2}{k^2\sigma^2} f(y)\,dy + \int_{\mu+k\sigma}^{\infty} \frac{(y-\mu)^2}{k^2\sigma^2} f(y)\,dy$$

$$\leq \frac{1}{k^2}\frac{1}{\sigma^2}\int_{-\infty}^{\infty} (y-\mu)^2 f(y)\,dy = \frac{1}{k^2}$$

Normal Probability Distribution (Gaussian)

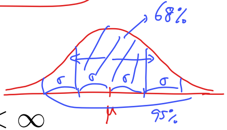
$N(3, 5)$

Stdev $= \sqrt{5}$

mean

variance

$Y \sim N(\mu, \sigma^2)$ iff it has density



$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/(2\sigma^2)}, \quad -\infty < y < \infty$$

- $f(\mu - t) = f(\mu + t) \quad \forall t$
- $f''(x) = 0 \Rightarrow x = \mu \pm \sigma$
- $f'(\mu) = 0$

for $\sigma > 0$ and $-\infty < \mu < \infty$.

$E(Y) = \mu$ and $V(Y) = \sigma^2$.

- $\displaystyle\int_{-\infty}^{\infty} y \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) dy = \mu$
- $\displaystyle\int_{-\infty}^{\infty} y^2 f(y)\, dy = \sigma^2 + \mu^2$

$\displaystyle\int f(y)\, dy = 1$

$\displaystyle\int e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \sqrt{2\pi}\,\sigma$

$$\bullet \quad \int_{-\infty}^{\infty} y^2 e^{-y^2}\, dy = \underbrace{\int_{\infty}^{\infty} y^2 \cdot \frac{1}{\sqrt{2\pi}\cdot\left(\frac{1}{\sqrt{2}}\right)} e^{-y^2}\, dy}_{} \cdot \sqrt{2\pi}\left(\frac{1}{\sqrt{2}}\right)$$

$$y^2 \to \frac{(y-0)^2}{2\cdot\left(\frac{1}{\sqrt{2}}\right)^2}$$

$\Downarrow$

$E(Y^2)$, when $Y \sim N\left(0, \left(\frac{1}{\sqrt{2}}\right)^2\right)$

$\parallel$

$\sigma^2 + \mu^2 = \frac{1}{2}$

$$= \frac{1}{2} \cdot \sqrt{2\pi} \cdot \frac{1}{\sqrt{2}} = \frac{\sqrt{\pi}}{2}$$

$$\int_{-\infty}^{\infty} (y-1)^2 e^{-(y^2+2y)}\, dy =$$

$$I = \int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2}\, dy = \sqrt{2\pi}$$

$$\cdot \int_{-\infty}^{\infty} e^{-y^2}\, dy = \sqrt{\pi}$$

$\longrightarrow \int$ Normal density $= 1$

Standard Normal Probability Distribution

If $Y \sim N(\mu, \sigma^2)$, then the standardized r.v., $Z = (Y - \mu)/\sigma$ has the standard normal distribution, denoted as $Z \sim N(0, 1)$.

Using standard normal distribution, we can calculate probabilities of normal r.v. Y with different mean and standard deviation.

Example) The number of calories in a salad on the lunch menu is normally distributed with $\mu = 200$ and $\sigma = 5$. Find the probability that the salad you randomly select will contain.

(a) More than 208 calories $P(Y > 208) = P\left(Z > \frac{208-200}{5}\right) = P(Z > 1.6)$

$= .0548 = p$

(b) Between 190 and 200 calories

(c) Select 10 salads. P(at least 8 of them has calories > 208)

X: # salads out of 10 that has > 208. $\sim B(10, .0548) = 1 - \binom{10}{9} p^9 q^1 - \binom{10}{10} p^{10} q^0$

$Y \sim N(\mu, \sigma^2)$ $\qquad Z = \frac{Y-\mu}{\sigma} \sim N(0,1)$

- $E(aY+b) = aE(Y)+b$
- $V(aX+b) = a^2 V(X)$

True for any r.v.

① $Y \sim N(\mu, \sigma^2)$

then, $aY+b \sim N(a\mu+b,\ a^2\sigma^2)$ ☆

MGF of Y:

$$\int_{-\infty}^{\infty} e^{ty} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) dy$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma} \int \exp\left[-\frac{1}{2\sigma^2}\left(y^2 - 2\mu y + \mu^2 - 2\sigma^2 t y\right)\right] dy$$

→ extracting a form of Normal density

$$= \exp\left(\mu t + \tfrac{1}{2}\sigma^2 t^2\right) \underbrace{\int \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{1}{\underset{\text{variance}}{2\sigma^2}}\left(y - \underset{\text{mean}}{(\mu+\sigma^2 t)}\right)^2\right) dy}_{\int N(\mu+\sigma^2 t,\ \sigma^2)\ \text{density} \ = 1}$$

☆☆

$$\boxed{= \exp\left(\mu t + \tfrac{1}{2}\sigma^2 t^2\right)}$$

$$Z = \frac{Y-\mu}{\sigma} = \frac{1}{\sigma}Y - \frac{\mu}{\sigma} \quad (= aY + b)$$

$$m_Z(t) = e^{bt} \cdot m_Y(at) = e^{-\frac{\mu}{\sigma}t} \cdot \exp\left(\mu \cdot \frac{1}{\sigma}t + \frac{1}{2}\sigma^2 \cdot \frac{1}{\sigma^2}t^2\right)$$

$$= e^{\frac{1}{2}t^2} = \exp\left(0 \cdot t + \frac{1}{2} \cdot 1^2 \cdot t^2\right)$$

: mgf of $N(0, 1)$

Normal mgf $= \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$

$$= 1 + \left(\mu t + \frac{1}{2}\sigma^2 t^2\right) + \frac{1}{2!}\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)^2 + \frac{1}{3!}\left(\quad\right)^3 \cdots$$

$$= 1 + \underbrace{\mu}_{E(Y)} t + \frac{1}{2}t^2\left(\underbrace{\mu^2 + \sigma^2}_{E(Y^2)}\right) + \frac{1}{3!}t^3\left(\underbrace{3\mu\sigma^2 + \mu^3}_{E(Y^3)}\right) + \cdots$$

$Z \sim N(0,1)$ $\quad f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$: even ft.

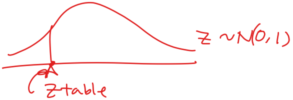
$$E(Z^k) = \int_{-\infty}^{\infty} z^k f(z)\, dz = 0 \quad \text{if } k \text{ is odd.}$$

Examples

Example) The raw scores in a national aptitude test are normally distributed with $\mu = 506$ and $\sigma = 81$.

(a) What proportion of the candidates scored below 574?

(b) Find the 30th percentile of the scores.



Example) According to government reports (USDHEW 79-1659), the heights of adult male residents of the United States are approximately normally distributed with a mean of 69.0 inches and a standard deviation of 2.8 inches. If a clothing manufacturer wants to limit his market to the central 80% of the adult male population, what range of heights should be targeted?