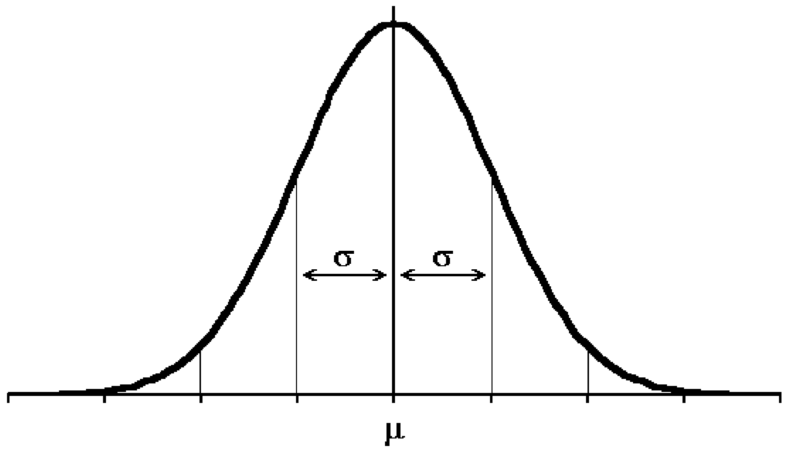
STAT 400 Review (4)

Fall 2020
A. Stepanov

Normal (Gaussian) Distribution.



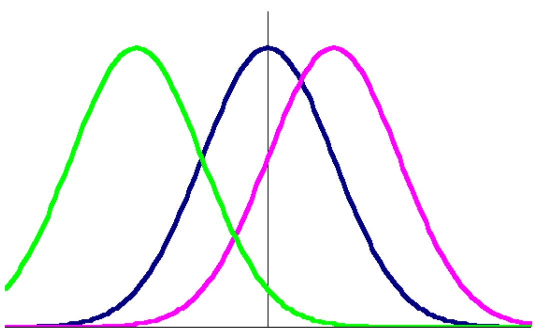
μ – mean

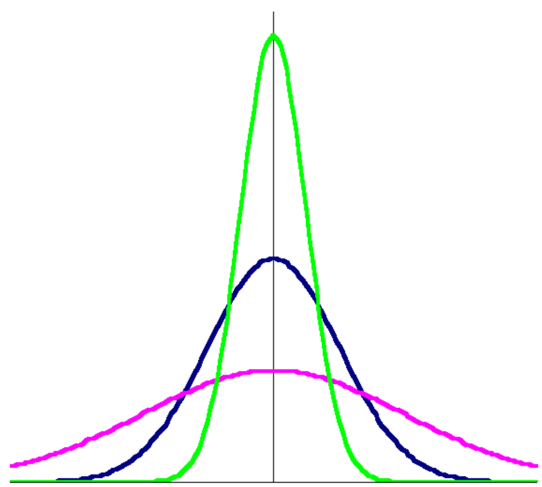
σ – standard deviation

$$\mathbf{N}\left(\mu, \sigma^2\right)$$

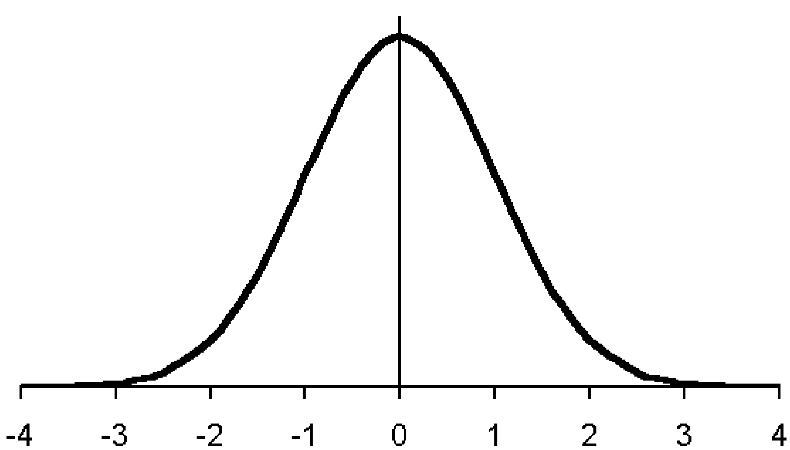
$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-(x-\mu)^2 / 2\sigma^2},$$

$$-\infty < x < \infty.$$





Standard Normal Distribution.



mean

0

standard deviation

1

$$\mathrm{N}(0, 1)$$

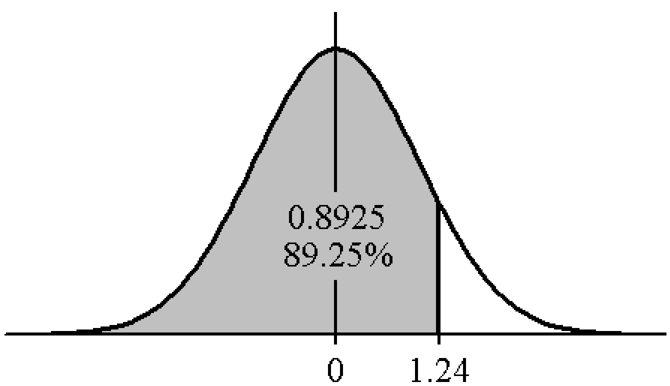
Example:

For the standard normal distribution, find the area to the left of

$z = 1.24$

1.24

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.0
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.53
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.57
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.61
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.65
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.68
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.72
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.75
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.78
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.81
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.83
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.86
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.88
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.90
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.91
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.93
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.94
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.95



Area to the left of $z = 1.24$ is **0.8925**.

$Z \sim N(0,1)$

$X \sim N(\mu, \sigma^2)$

$$Z = \frac{X - \mu}{\sigma}$$

$$X = \mu + \sigma Z$$

EXCEL:

$=$NORM.S.DIST(z)	gives	$\mathrm{P}(\mathrm{Z} \le z) = \Phi(z)$
$=$NORM.S.INV(p)	gives	z such that $\mathrm{P}(\mathrm{Z} \le z) = p$
$=$NORM.DIST$(x, \mu, \sigma, 1)$	gives	$\mathrm{P}(\mathrm{X} \le x)$, where X is $\mathrm{N}(\mu, \sigma^2)$
$=$NORM.DIST$(x, \mu, \sigma, 0)$	gives	$f(x)$, p.d.f. of $\mathrm{N}(\mu, \sigma^2)$
$=$NORM.INV(p, μ, σ)	gives	x such that $\mathrm{P}(\mathrm{X} \le x) = p$, where X is $\mathrm{N}(\mu, \sigma^2)$

R: If `mean` or `sd` are not specified they assume the default values of `0` and `1`, respectively.

`dnorm(x, mean = 0, sd = 1, log = FALSE)`

Density for the normal distribution with mean equal to `mean` and standard deviation equal to `sd`.

$f(x)$, p.d.f. of $\mathrm{N}(\mu, \sigma^2)$

`pnorm(q, mean = 0, sd = 1, lower.tail = TRUE, log.p = FALSE)`

Distribution function for the normal distribution with mean equal to `mean` and standard deviation equal to `sd`.

$\mathrm{P}(\mathrm{X} \le x)$, where X is $\mathrm{N}(\mu, \sigma^2)$

`qnorm(p, mean = 0, sd = 1, lower.tail = TRUE, log.p = FALSE)`

Quantile function and random generation for the normal distribution with mean equal to `mean` and standard deviation equal to `sd`.

x such that $\mathrm{P}(\mathrm{X} \le x) = p$, where X is $\mathrm{N}(\mu, \sigma^2)$

`rnorm(n, mean = 0, sd = 1)`

Random generation for the normal distribution with mean equal to `mean` and standard deviation equal to `sd`.

Random sample of size n from $\mathrm{N}(\mu, \sigma^2)$ distribution

1. Models of the pricing of stock options often make the assumption of a normal distribution. An analyst believes that the price of an *Initech* stock option varies from day to day according to normal distribution with mean \$9.22 and unknown standard deviation.

a) The analyst also believes that 77% of the time the price of the option is greater than \$7.00. Find the standard deviation of the price of the option.

$\mu = 9.22,\qquad \sigma = ?\qquad$ Know $P(X > 7.00) = 0.77.$

Find z such that $P(Z > z) = 0.77.$

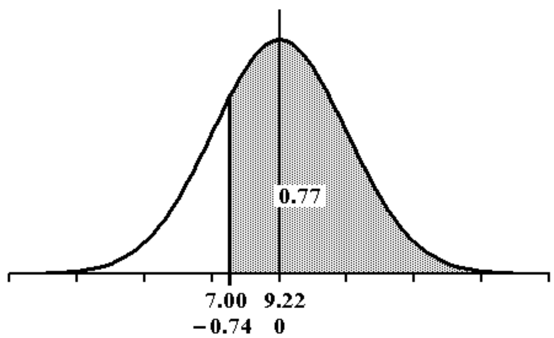
$\Phi(z) = 1 - 0.77 = \mathbf{0.23}.$

$z = \mathbf{-0.74}.$

$x = \mu + \sigma \cdot z.$

$7.00 = 9.22 + \sigma \cdot (-0.74).$

$\sigma = \mathbf{\$3.00}.$



b) Find the proportion of days when the price of the option is greater than \$10.00?

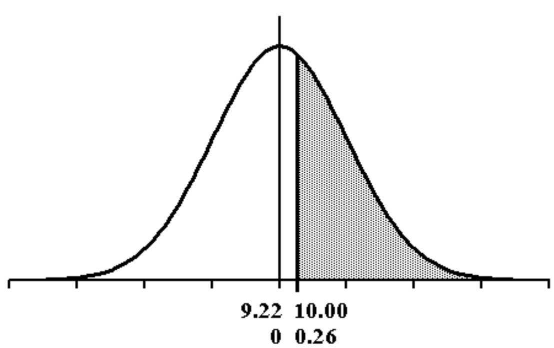
$$P(X > 10.00) = P\left(Z > \frac{10.00 - 9.22}{3.00}\right)$$

$$= P(Z > 0.26)$$

$$= 1 - \Phi(0.26)$$

$$= 1 - 0.6026$$

$$= \mathbf{0.3974}.$$



c) Following the famous "buy low, sell high" principle, the analyst recommends buying *Initech* stock option if the price falls into the lowest 14% of the price distribution, and selling if the price rises into the highest 9% of the distribution. Mr. Statman doesn't know much about history, doesn't know much about biology, doesn't know much about statistics, but he does want to be rich someday. Help Mr. Statman find the price below which he should buy *Initech* stock option and the price above which he should sell.

Need $x = ?$ such that $P(X < x) = 0.14$.

Find z such that $P(Z < z) = 0.14$.

The area to the left is $\mathbf{0.14} = \Phi(z)$.

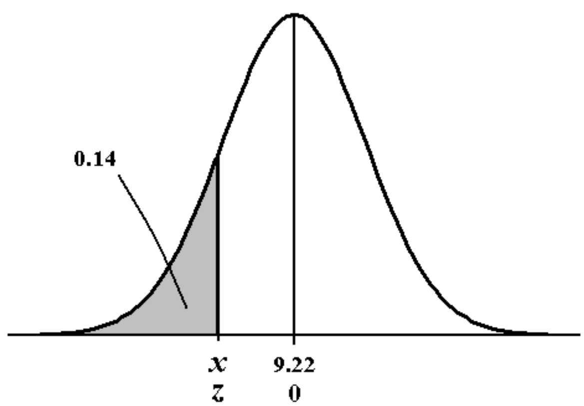
$z = \mathbf{-1.08}$.

$x = \mu + \sigma \cdot z$.

$x = 9.22 + 3 \cdot (-1.08)$

$= \mathbf{\$5.98}$.

Buy if the price is below \$5.98.



Need $x = ?$ such that $P(X > x) = 0.09$.

① Find z such that $P(Z > z) = 0.09$.

The area to the left is $\mathbf{0.91} = \Phi(z)$.

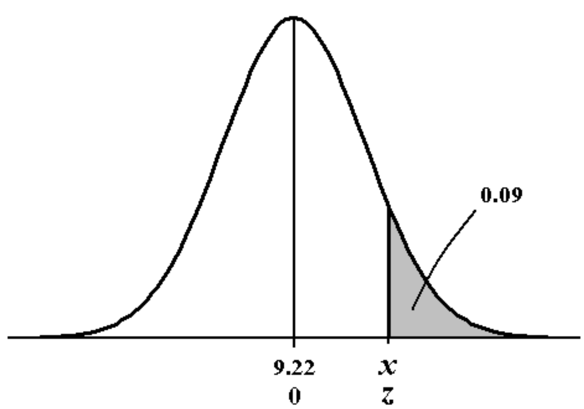
$z = \mathbf{1.34}$.

② $x = \mu + \sigma \cdot z$.

$x = 9.22 + 3 \cdot (1.34)$

$= \mathbf{\$13.24}$.

Sell if the price is above \$13.24.



$X \sim N(\mu, \sigma^2) \qquad \Leftrightarrow \qquad M_X(t) = e^{\mu t + \sigma^2 t^2/2}.$

Proof: $$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} \cdot \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2} dx$$

$$= \int_{-\infty}^{\infty} e^{t(\mu + \sigma z)} \cdot \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

$$= e^{\mu t + \sigma^2 t^2/2} \cdot \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(z - \sigma t)^2/2} dz$$

$$= e^{\mu t + \sigma^2 t^2/2},$$

since $\frac{1}{\sqrt{2\pi}} e^{-(z-\sigma t)^2/2}$ is the probability density function of a $N(\sigma t, 1)$ random variable.

Let $Y = aX + b$. Then $M_Y(t) = e^{bt} M_X(at)$.

If X is normally distributed with mean μ and standard deviation σ, Y is normally distributed with mean $a\mu + b$ and variance $a^2\sigma^2$ (standard deviation $|a|\sigma$).

2. Suppose the average daily temperature [in degrees Fahrenheit] in July in Anytown is a random variable T with mean $\mu_T = 85$ and standard deviation $\sigma_T = 7$. The daily air conditioning cost Q, in dollars, for Anytown State University, is related to T by

$$Q = 120\,T + 750.$$

Suppose that T is a normal random variable. Compute the probability that the daily air conditioning cost on a typical July day for the factory will exceed \$12,210.

Q has Normal distribution.

$\mu_Q = 120\,\mu_T + 750 = 120 \cdot 85 + 750 = \$10{,}950.$

$\sigma_Q^2 = (120)^2 \cdot \sigma_T^2 = (120)^2 \cdot 7^2 = 840^2. \qquad \sigma_Q = \$840.$

$$P(Q > 12{,}210) = P\left(Z > \frac{12{,}210 - 10{,}950}{840}\right)$$

$$= P(Z > 1.50) = 1 - \Phi(1.50)$$

$$= 1 - 0.9332 = \mathbf{0.0668}.$$

OR

$12{,}210 = 120\,T + 750. \qquad \Leftrightarrow \qquad T = 95.5.$

$$P(Q > 12{,}210) = P(T > 95.5)$$

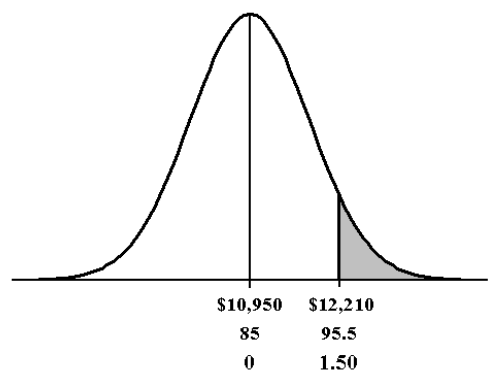
$$= P\left(Z > \frac{95.5 - 85}{7}\right)$$

$$= P(Z > 1.50)$$

$$= 1 - \Phi(1.50)$$

$$= 1 - 0.9332$$

$$= \mathbf{0.0668}.$$



Recall: $\quad \mathrm{E}(a\mathrm{X} + b\mathrm{Y}) = a\,\mathrm{E}(\mathrm{X}) + b\,\mathrm{E}(\mathrm{Y}),$

$\mathrm{Var}(a\mathrm{X} + b\mathrm{Y}) = a^2\,\mathrm{Var}(\mathrm{X}) + 2\,a\,b\,\mathrm{Cov}(\mathrm{X}, \mathrm{Y}) + b^2\,\mathrm{Var}(\mathrm{Y}).$

If $X_1, X_2, \ldots, X_n$ are n random variables and $a_0, a_1, a_2, \ldots, a_n$ are $n+1$ constants, then the random variable $U = a_0 + a_1 X_1 + a_2 X_2 + \ldots + a_n X_n$ has mean

$$\mathrm{E}(U) = a_0 + a_1\,\mathrm{E}(X_1) + a_2\,\mathrm{E}(X_2) + \ldots + a_n\,\mathrm{E}(X_n)$$

and variance

$$\mathrm{Var}(U) = \sum_{i=1}^{n} a_i^2\,\mathrm{Var}(X_i) + \sum_{i=1}^{n}\sum_{\substack{j=1 \\ i \neq j}}^{n} a_i\,a_j\,\mathrm{Cov}(X_i, X_j)$$

$$= \sum_{i=1}^{n} a_i^2\,\mathrm{Var}(X_i) + 2\sum_{i<j}\sum a_i\,a_j\,\mathrm{Cov}(X_i, X_j)$$

If $X_1, X_2, \dots, X_n$ are n independent random variables and $a_0, a_1, a_2, \dots, a_n$ are $n+1$ constants, then the random variable $U = a_0 + a_1 X_1 + a_2 X_2 + \dots + a_n X_n$ has variance

$$\mathrm{Var}(U) = a_1^2\,\mathrm{Var}(X_1) + a_2^2\,\mathrm{Var}(X_2) + \dots + a_n^2\,\mathrm{Var}(X_n)$$

Also

$$M_U(t) = e^{a_0 t} \cdot M_{X_1}(a_1 t) \cdot M_{X_2}(a_2 t) \cdot \dots \cdot M_{X_n}(a_n t).$$

If $X_1, X_2, \dots, X_n$ are normally distributed random variables, then U is also normally distributed.

3. Models of the pricing of stock options often make the assumption of a normal distribution. An investor believes that the price of an *Burger Queen* stock option is a normally distributed random variable with mean \$18 and standard deviation \$3. He also believes that the price of an *Dairy King* stock option is a normally distributed random variable with mean \$14 and standard deviation \$2. Assume the stock options of these two companies are independent. The investor buys 8 shares of *Burger Queen* stock option and 9 shares of *Dairy King* stock option. What is the probability that the value of this portfolio will exceed \$300?

BQ has Normal distribution, $\mu_{BQ} = \$18$, $\sigma_{BQ} = \$3$.

DK has Normal distribution, $\mu_{DK} = \$14$, $\sigma_{DK} = \$2$.

Value of the portfolio $VP = 8 \times BQ + 9 \times DK$.

Then VP has Normal distribution.

$$\mu_{VP} = 8 \times \mu_{BQ} + 9 \times \mu_{DK} = 8 \times 18 + 9 \times 14 = \$270.$$

$$\sigma_{VP}^2 = 8^2 \times \sigma_{BQ}^2 + 9^2 \times \sigma_{DK}^2 = 64 \times 9 + 81 \times 4 = 900. \qquad \sigma_{VP} = \$30.$$

$$P(VP > 300) = P\left(Z > \frac{300 - 270}{30}\right) = P(Z > 1.00) = 1 - 0.8413 = \mathbf{0.1587}.$$

4. In Neverland, the weights of adult men are normally distributed with mean of 170 pounds and standard deviation of 10 pounds, and the weights of adult women are normally distributed with mean of 125 pounds and standard deviation of 8 pounds. Six women and four men got on an elevator. Assume that all their weights are independent. What is the probability that their total weight exceeds 1500 pounds?

$$\text{Total} = W_1 + W_2 + W_3 + W_4 + W_5 + W_6 + M_1 + M_2 + M_3 + M_4.$$

$$\begin{aligned} E(\text{Total}) &= E(W_1) + E(W_2) + E(W_3) + E(W_4) + E(W_5) + E(W_6) \\ &\quad + E(M_1) + E(M_2) + E(M_3) + E(M_4) \\ &= 125 + 125 + 125 + 125 + 125 + 125 + 170 + 170 + 170 + 170 = 1430. \end{aligned}$$

$$\begin{aligned} \text{Var}(\text{Total}) &= \text{Var}(W_1) + \text{Var}(W_2) + \text{Var}(W_3) + \text{Var}(W_4) + \text{Var}(W_5) + \text{Var}(W_6) \\ &\quad + \text{Var}(M_1) + \text{Var}(M_2) + \text{Var}(M_3) + \text{Var}(M_4) \\ &= 64 + 64 + 64 + 64 + 64 + 64 + 100 + 100 + 100 + 100 = 784. \end{aligned}$$

$$\text{SD}(\text{Total}) = \sqrt{784} = 28.$$

Total has Normal distribution.

$$P(\text{ Total} > 1500 \text{ }) = P\left(Z > \frac{1500 - 1430}{28} \right) = P(\text{ } Z > 2.50 \text{ }) = 1 - 0.9938 = \mathbf{0.0062}.$$

<u>Note</u>: It is tempting to set Total = 6 W + 4 M, but that would imply that the six women who got on the elevator all have the same weight, and so do the four men, which is most likely not the case here.

5. A machine fastens plastic screw-on caps onto containers of motor oil. If the machine applies more torque than the cap can withstand, the cap will break. Both the torque applied and the strength of the caps vary. The capping machine torque has the normal distribution with mean 7.9 inch-pounds and standard deviation 0.9 inch-pounds. The cap strength (the torque that would break the cap) has the normal distribution with mean 10 inch-pounds and standard deviation 1.2 inch-pounds. The cap strength and the torque applied by the machine are independent.

a) What is the probability that a cap will break while being fastened by the capping machine? That is, find the probability $\mathrm{P}(\text{Strength} < \text{Torque})$.

Need $\mathrm{P}(\text{Strength} < \text{Torque}) = \mathrm{P}(\text{Strength} - \text{Torque} < 0) = ?$

$\mathrm{E}(\text{Strength} - \text{Torque}) = 10 - 7.9 = 2.1.$

$\mathrm{Var}(\text{Strength} - \text{Torque}) = (1)^2 \cdot 1.2^2 + (-1)^2 \cdot 0.9^2 = 2.25.$

$\mathrm{SD}(\text{Strength} - \text{Torque}) = 1.5.$

$(\text{Strength} - \text{Torque})$ is normally distributed.

$$\mathrm{P}(\text{Strength} - \text{Torque} < 0) = \mathrm{P}\left(Z < \frac{0 - 2.1}{1.5}\right) = \mathrm{P}(Z < -1.40) = \mathbf{0.0808}.$$

b) Suppose the mean torque of the capping machine can be adjusted without changing its standard deviation. What should the new mean torque be so that the machine would break no more than 2.5% of caps?

Want $\mathrm{P}(\text{Strength} < \text{Torque}) = \mathrm{P}(\text{Strength} - \text{Torque} < 0) = 0.0250.$

$\mathrm{E}(\text{Strength} - \text{Torque}) = 10 - \mu.$

$\mathrm{Var}(\text{Strength} - \text{Torque}) = (1)^2\, 1.2^2 + (-1)^2\, 0.9^2 = 2.25.$

$\mathrm{SD}(\text{Strength} - \text{Torque}) = 1.5.$

$(\text{Strength} - \text{Torque})$ is normally distributed.

$$P(\text{Strength} - \text{Torque} < 0) = P\left(Z < \frac{0 - 10 + \mu}{1.5}\right) = P\left(Z < \frac{\mu - 10}{1.5}\right).$$

$P(Z < -1.96) = 0.0250.$

$\frac{\mu - 10}{1.5} = -1.96.$ $\mu = \mathbf{7.06}.$

6. One piece of PVC pipe is to be inserted inside another piece. The length of the first piece is normally distributed with mean value 20 in. and standard deviation 0.7 in. The length of the second piece is a normal random variable with mean and standard deviation 15 in. and 0.6 in., respectively. The amount of overlap is normally distributed with mean value 1 in. and standard deviation 0.2 in. Assuming that the lengths and amount of overlap are independent of one another, what is the probability that the total length after insertion is between 32.65 in. and 35.35 in.?

Total = First + Second – Overlap.

E(Total) = E(First) + E(Second) – E(Overlap) = 20 + 15 – 1 = 34.

Var(Total) = Var(First) + Var(Second) + $(-1)^2$ Var(Overlap)

$= 0.7^2 + 0.6^2 + 0.2^2 = 0.49 + 0.36 + 0.04 = 0.89.$

SD(Total) = $\sqrt{0.89} = 0.9434.$ Total has Normal distribution.

$$P(32.65 < \text{Total} < 35.35) = P\left(\frac{32.65 - 34}{0.9434} < Z < \frac{35.35 - 34}{0.9434}\right) = P(-1.43 < Z < 1.43)$$

$$= 0.9236 - 0.0764 = \mathbf{0.8472}.$$

7. At Sam's Butcher Shop, ground beef packages vary in weight according to a normal distribution with a mean of 3.1 pounds and a standard deviation of 0.2 pounds, and are sold for \$1 per pound. Packages of Bratwurst vary in weight according to a normal distribution with a mean of 2.6 pounds and a standard deviation of 0.3 pounds, and are sold for \$3 per pound. Anticipating nice weather during the weekend, Dick buys two packages of ground beef and one package of Bratwurst, selecting the packages at random. What is the probability that Dick would exceed the \$15 limit "suggested" by his wife Jane? (Assume independence.)

$$\text{Total} = \$3 \cdot \text{Bratwurst} + \$1 \cdot \text{GrBeef}_1 + \$1 \cdot \text{GrBeef}_2.$$

$$E(\text{Total}) = \$3 \cdot E(\text{Bratwurst}) + \$1 \cdot E(\text{GrBeef}_1) + \$1 \cdot E(\text{GrBeef}_2)$$

$$= \$3 \cdot 2.6 + \$1 \cdot 3.1 + \$1 \cdot 3.1 = \$14.$$

$$\text{Var}(\text{Total}) = 3^2 \cdot \text{Var}(\text{Bratwurst}) + 1^2 \cdot \text{Var}(\text{GrBeef}_1) + 1^2 \cdot \text{Var}(\text{GrBeef}_2)$$

$$= 3^2 \cdot 0.3^2 + 1^2 \cdot 0.2^2 + 1^2 \cdot 0.2^2 = 0.89.$$

$\text{SD}(\text{Total}) = \sqrt{0.89} = \$0.9434.$ Total has Normal distribution.

$$P(\text{Total} > 15) = P\left(Z > \frac{15 - 14}{0.9434}\right) = P(Z > 1.06) = 1 - 0.8554 = \mathbf{0.1446}.$$

8. Show that the odd moments of $N(0, \sigma^2)$ are zero and the even moments are

$$\mu_{2n} = \frac{(2n)!\sigma^{2n}}{2^n (n)!}$$

Taylor Formula:

$$M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!} M^{(r)}(0) = \sum_{r=0}^{\infty} \frac{t^r}{r!} E\left(X^r\right) = \sum_{r=0}^{\infty} \frac{t^r}{r!} \mu_r .$$

Since X is $N(0, \sigma^2)$,

$$M_X(t) = \exp\left\{\frac{\sigma^2 t^2}{2}\right\} = \sum_{n=0}^{\infty} \frac{\sigma^{2n} t^{2n}}{2^n n!}$$

Therefore,

if r is odd, $\quad \mu_r = 0,$

if $r = 2n$ is even, $\quad \dfrac{\sigma^{2n}}{2^n n!} = \dfrac{1}{r!} \mu_r \quad \Rightarrow \quad \mu_{2n} = \dfrac{(2n)!\sigma^{2n}}{2^n (n)!}.$

OR

Def $\quad \Gamma(x) = \int_0^{\infty} u^{x-1} e^{-u} \, du, \quad x > 0$

$$\Gamma(x) = (x-1)\,\Gamma(x-1), \quad x > 1$$

$\Gamma(n) = (n-1)!$ if n is a positive integer

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

$$\mu_{2n} = \int_{-\infty}^{\infty} x^{2n} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-x^2/2\sigma^2}\, dx = \int_{0}^{\infty} x^{2n} \frac{2}{\sqrt{2\pi}\,\sigma} e^{-x^2/2\sigma^2}\, dx = \ldots$$

$$u = \frac{x^2}{2\sigma^2} \qquad du = \frac{x\,dx}{\sigma^2} \qquad dx = \frac{du\,\sigma}{\sqrt{2u}}$$

$$\ldots = \int_{0}^{\infty} 2^n \sigma^{2n} \frac{1}{\sqrt{\pi}} u^{n-1/2} e^{-u}\, du = 2^n \sigma^{2n} \frac{1}{\sqrt{\pi}} \Gamma\left(n+\frac{1}{2}\right).$$

$$\Gamma\left(n+\frac{1}{2}\right) = \left(n-\frac{1}{2}\right)\cdot\left(n-\frac{3}{2}\right)\cdot\left(n-\frac{5}{2}\right)\cdot\ldots\cdot\frac{3}{2}\cdot\frac{1}{2}\cdot\sqrt{\pi}$$

$$= \frac{1}{2^n}\cdot(2n-1)\cdot(2n-3)\cdot(2n-5)\cdot\ldots\cdot 3\cdot 1\cdot\sqrt{\pi}$$

$$= \frac{1}{2^n}\cdot\frac{(2n)!}{(2n)\cdot(2n-2)\cdot(2n-4)\cdot\ldots\cdot 4\cdot 2}\cdot\sqrt{\pi}$$

$$= \frac{1}{2^n}\cdot\frac{(2n)!}{2^n\cdot(n)!}\cdot\sqrt{\pi}$$

$$\Rightarrow \qquad \mu_{2n} = \frac{(2n)!\,\sigma^{2n}}{2^n\,(n)!}.$$