② $E(\underbrace{X(X-1)(X-2)\cdots(X-(n-1))}_{\text{mulitiplcation of } N \text{ terms.}})$

$$= \left[\frac{d^n}{dz^n} G_X(z)\right]\Bigg|_{z=1}$$

Ex: Suppose we know that for a Poisson R.V with para α

$G_X(z) = e^{\alpha(z-1)}$

Q: $E(X)$, $E(X(X-1))$, $E(X^2)$, $Var(X)$?

Ans: $E(X) = \frac{d}{dz} G_X(z)\Big|_{z=1} = \alpha e^{\alpha(z-1)}\Big|_{z=1}$

$= \alpha$

$E(X(X-1)) = \left[\frac{d^2}{dz^2} G_X(z)\right]\Big|_{z=1} = \alpha^2 e^{\alpha(z-1)}\Big|_{z=1}$

$= \alpha^2$

$E(X^2) = E(X(X-1)) + E(X) = \alpha + \alpha^2$

$Var(X) = E(X^2) - \alpha^2 = (\alpha + \alpha^2) - \alpha^2$

$= \alpha$

Functions of R.Vs.

Recall the advantage of considering R.V.s.
is ① we can compute the weighted average
② we can easily generate new R.V from an old R.V.

Ex X is a R.V.

$Y_1 = X^2$ is another R.V.

$Y_2 = \frac{1}{X^2-1}$ is another R.V

Basically for any function $f(x)$

$Y = f(X)$ is a new R.V.

How to describe the new W.A of Y?

Ans: Method I: The most universal method is to ① compute the cdf of Y first

$$F_Y(y) = P(Y \leq y) = P(f(X) \leq y)$$

② Use cdf to obtain pmf/pdf

Method 2: Sometimes we can compute the pmf / pdf of Y directly

* In the past, we have discussed different types of functions ~~&~~ the W.A of $Y = f(X)$.

I.e. $X \longrightarrow$ System / function $\xrightarrow{Y}$

Knowing the W.A of X, $\Rightarrow$ the W.A of Y

Ex: f is a "quantizer" HW6 Q2.

$f(0, x) = \max(x, 0)$ — half-wave rectifier

Other important functions:

$f(x) = |x|$ — full-wave rectifier

$f(x) = \min(x, 10)$ — limiter / clipper

* You need some practice on computing the W.A of Y from X.

* The most important function is the "linear functions"

$$Y = aX + b.$$

Ex: X is a geometric R.V with p.

$$Y = 2X + 1$$

Find the pmf of Y.

Ans: $P_k = P(Y = k)$

$$= P\left(X = \frac{k-1}{2}\right)$$

$$= \begin{cases} 0 & \text{if } k \text{ is even} \\ 0 & \text{if } \frac{k-1}{2} < 0 \\ p(1-p)^{\frac{k-1}{2}} & \text{if } k \text{ is odd } \& \ \frac{k-1}{2} \geq 0 \end{cases}$$

Basically the position of the prob mass has to be relocated.

for conti R.V. with $Y = aX + b$

We have a quick formula

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$

if $a \neq 0$.

(see Example 4.31 for detailed derivation)

Q: If $a = 0$, what is the pdf of Y?

Ans: $f_Y(y) = 1 \cdot \delta(y-b)$ #.

Note Title 3/4/2011

* Expectation & Variance of $Y = aX + b$.

For any R.V. X (conti / discrete / mixed type)

We have $E(Y) = aE(X) + b$

$Var(Y) = |a|^2 Var(X)$

the center of the W.A

the expected "$\boxed{\text{squared}}$ distance" to the center

Ex: Continue from the example that X: geometric. $Y = 2X + 1$

Q: $E(Y)$, $Var(Y) = ?$ Ans: $E(Y) = 2E(X) + 1 = 2\frac{1}{p} + 1$, $Var(Y) = 4 \times \frac{1-p}{p^2}$

Q: Is Y a geometric R.V.?

Ans: No.

* Linear functions of Gaussian R.V.

Theorem: If X is a Gaussian R.V with μ_X, σ_X^2, and $Y = aX + b$.

then Y is also a Gaussian R.V with $\mu_Y = a\mu_X + b$

$\sigma_Y^2 = |a|^2 \sigma_X^2$

Therefore, whenever we see a Gsn R.V. X with mean μ & variance σ^2, we should view it as a linear function of

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

where Z is a Gsn with $\mu = 0$ (zero mean) & $\sigma^2 = 1$ (unit variance).

Such Z with zero-mean & unit variance is called the $\boxed{\text{standard}}$ Gsn R.V.

Ex: X is a Gaussian R.V with μ, σ^2

Find $P(X \leq 3)$ in terms of the prob of a standard Gsn R.V. Z

Ans: $P(X \leq 3)$

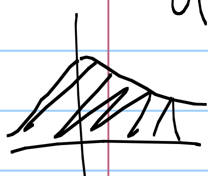
$$= P(\mu + \sigma Z \leq 3)$$

$$= P\left(Z \leq \frac{3-\mu}{\sigma}\right)$$

Q: Why are we interested in a Standard Gsn?

Ans: We can construct a standard table for the cdf of Z. No need to construct multiple tables for different μ, σ^2 values.

In practice, we can compute a table of the cdf of Z.



$$\Phi_Z(z) = \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi 1^2}} e^{-\frac{(x-0)^2}{2\cdot 1}} \, dx$$

$F_Z(z)$; Z is often omitted. (Don't be confused with the characteristic functions.)

Continue from the previous example,

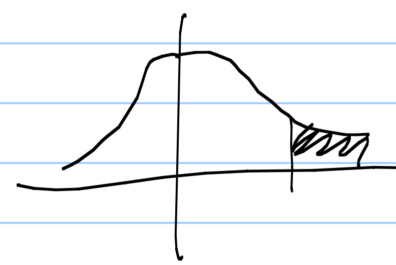
Then $P(X \leq 3)$ is obtained by looking up the value of $\Phi\left(\frac{3-\mu}{\sigma}\right)$ ※

Other "tables" for computing the prob of a Gssn R.V.

① Q functions: very popular in ECE.

$$Q(x) = 1 - \Phi(x)$$

$$= P(Z > x)$$



for the previous example

$$P(X \leq 3) = 1 - Q\left(\frac{3-\mu}{\sigma}\right)$$ P.169

Other related "tables" $\boxed{\text{erf}}$ $\boxed{\text{erfc}}$ in statistics.

Ex: X is a Gsn with $\mu = 2$ $\sigma^2 = 4$.

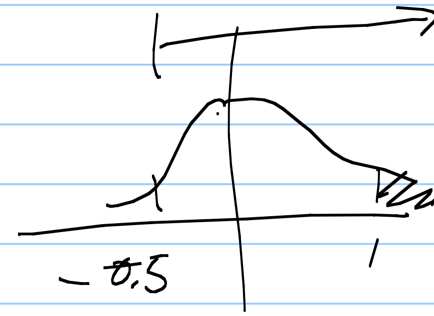
Q Find the prob $P(1 \le X \le 4)$ in terms of the Q function.

Ans: $P(1 \le X \le 4)$

$= P(1 \le 2 + 2Z \le 4)$

$= P(-0.5 \le Z \le 1)$

$= Q(-0.5) - Q(1)$



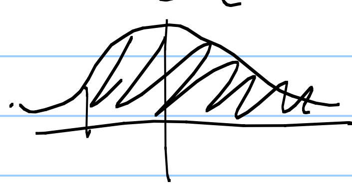
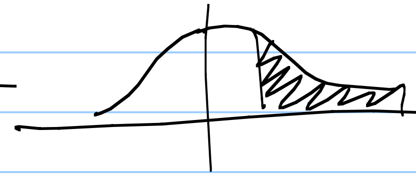
Q: You only know

$Q(0.5) = 0.309$ (Table 4.2)
$Q(1) = 0.159$

Find the value of $P(1 \le X \le 4)$

Ans: the standard Gsn variable Z is symmetric

$\Rightarrow Q(-0.5) = 1 - Q(0.5)$

 $= 1 -$ 

$\Rightarrow P(1 \le X \le 4) = (1 - Q(0.5)) - Q(1)$
$= 0.532$

In MATLAB, there is no Q function implemented. You can create the following .m file. fo

```
function y = q_ece(x)
y=0.5*erfc(x/sqrt(2));
```

Ex: X is a Gaussian with $\mu=1$, $\sigma^2=4$.

$Y=2X+1$. Find the prob

$P(1\le Y\le 2 \text{ or } 3\le Y\le 4)$

Ans: Y has mean $2\times 1+1=3$

variance $2^2\times 4=16=4^2$

$\Rightarrow Y=3+4Z$

$P(1\le Y\le 2 \text{ or } 3\le Y\le 4)$

$=P(1\le 3+4Z\le 2,\ 3\le 3+4Z\le 4)$

$=P(\frac{-2}{4}\le Z\le\frac{-1}{4},\ 0\le Z\le\frac{1}{4})$

$=Q(\frac{-2}{4})-Q(\frac{-1}{4})+Q(0)-Q(\frac{1}{4})$

$=(1-Q(\frac{1}{2}))-(1-Q(\frac{1}{4}))+Q(0)-Q(\frac{1}{4})$

$=Q(0)-Q(\frac{1}{2})=\frac{1}{2}-Q(\frac{1}{2})$ #

Note Title 3/2/2011

= Summary

* For any R.V. X. & $Y = aX + b$.

$\mu_Y = a\mu_X + b$, $Var(Y) = a^2 Var(X)$

* Gaussian R.V + linear transformation

Two important conclusions.

① If X is a Gsn with μ_X, σ_X^2 and $Y = aX + b$. then Y is also a Gsn with

$$\mu_Y = a\mu_X + b$$
$$\sigma_Y^2 = a^2 \sigma_X^2$$

② If X is a Gsn with μ_X, σ_X^2 then X can be viewed as a generated by $X = \mu_X + \sigma_X Z$ where

Z is a mean 0, variance 1 Gsn R.V, called the standard Gsn.

③ Computing the prob of Z is achieved by table look-up. $F_Z(z) \triangleq \Phi(z) = P(Z \le z)$

or $Q(x) = P(Z > x) = 1 - \Phi(z)$

So far, we assume that we know the W.A completely (say pdf/pmf/cdf/charact... moment generating/prob

But in many cases we only know part of the W.A. Can we still do some meaningful implication?

Probability Bounds:

① Union bounds.

If we know ① $P(0<X<3)$ and ② $P(1<X<5)$.

Q: what is the estimate of

$$P(0<X<3 \text{ or } 1<X<5)$$

Ans: $\max(P(0<X<3), P(1<X<5)) \le \; ? \; \le P(0<X<3) + P(1<X<5)$

② The Markov Inequality

We know ① $P(X<0)=0$

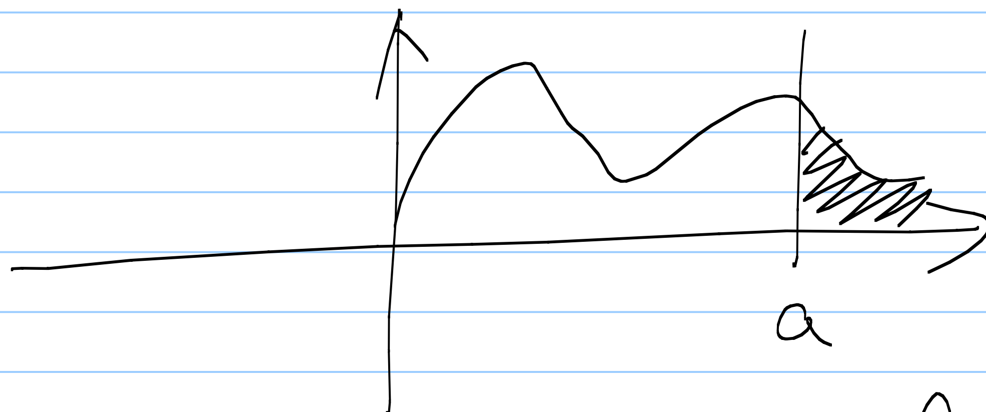
② $E(X)=m$ say $m=5$

We want to estimate $P(X \ge a)$ say $a=3$.

Specifically, what is the ~~largest~~ possible value of $P(X \geq a)$ while keeping the same mean value $E(X) = m$

say $a = 3$ say $m = 5$

Ans: $\boxed{P(X \geq a) \leq \frac{E(X)}{a}}$



pf: suppose $P(X \geq a)$ is fixed to K

how to design a W.A with the smallest E.

Ans: Choose $P(X=0) = 1-K$

$P(X=a) = K.$

the smallest possible $E(X)$ is $0 \times (1-K) + aK$.

$\Rightarrow$ For any other W.A.

$$E(X) \geq a \cdot P(X \geq a)$$

Note Title 3/4/2011

③ Chebyshev Inequality: A refinement of the Markov Inequality, which gives you more accurate estimate at the price of requiring more info. ② m ③ σ^2 (but does not need X to be positive)

$$P(|X-m| \geq a) \leq \frac{\sigma^2}{a^2}$$

pf: let $Y = (X-m)^2$

$$\Rightarrow P(|X-m| \geq a) = P(Y \geq a^2)$$

$$\leq \frac{E(Y)}{a^2}$$ by Markov inequality

$$\because E(Y) = E((X-m)^2) = Var(X)$$

$$= \frac{Var(X)}{\sigma^2}$$

Ex: I am recording the values of my stock portfolio, which has 100 stocks.

One day my computer crashes, I only remember the average price is 50.

Q: At most how many stocks can have the price ≥ 90?

Ans: By Markov inequality

$$P(X \geq 90) \leq \frac{50}{90}$$

$\Rightarrow$ At most $100 \times \frac{50}{90}$ stocks can have value larger than 90.

Q: Suppose I also remember the standard deviation is 20.
What is the maximum number of stocks having value ≥ 90.

Ans: $P(X \geq 90) = P(X-50 \geq 40)$

$$\leq P(|X-50| \geq 40)$$

$$\leq \frac{20^2}{40^2} = \frac{1}{4} = 0.25$$

Q: Suppose I also know it is a bell-shaped distribution
What is $P(X \geq 90)$?

Ans: X is Gsn with $\mu = 50$, $\sigma = 20$

$\Rightarrow X = 50 + 20Z$ where Z is standard Gsn

$$P(X > 90) = P(50 + 20Z > 90)$$

$$= P\left(Z > \frac{90-50}{20}\right)$$

$$= P(Z > 2) = Q(2) = 0.0228$$

Note Title 3/7/2011

(4) Chernoff Inequality: A further refinement of the Markov inequality

We need only (4) The moment generating function $X^*(s)$

Chernoff Inequality.

For any a

$$P(x \geq a) \leq e^{sa} X^*(s)$$

for any negative $s \leq 0$.

To have the tightest bound

$$\leq \min_{s \leq 0} e^{sa} X^*(s)$$

Take the min over all possible $s \leq 0$.

* Continue from our example. ~~For~~ a Gsn w. $\mu = 50$ $\sigma = 20$, $a = 90$

$$X^*(s) = e^{-s \cdot \mu + \frac{\sigma^2 s^2}{2}}$$

$$= e^{-50s + 200s^2}$$

$$e^{sa} \cdot X^*(s) = e^{90s - 50s + 200s^2} = e^{40s + 200s^2}$$

$$= e^{200(s+\frac{1}{10})^2 - 2}$$

$\Rightarrow$ The min value is when $s = -\frac{1}{10}$

$$\Rightarrow P(X \geq 90) < e^{-2} \approx 0.135$$

Pf ~~of~~ of the Chernoff bound. for any $s \leq 0$

$$P(X \geq a) = P(-sX \geq -sa)$$

$$= P(e^{-sX} \geq e^{-sa}) \leq \frac{E(e^{-sX})}{e^{-sa}} = e^{sa} X^*(s)$$

* Chernoff bound is tricky.

Since once knowing $X^*(s)$, we already know the exact distribution of X. We can thus use summation/integration to find the "exact" value of $P(X \geq a)$. However, finding the exact value of $P(X \geq a)$ is computationally intensive. In many cases, find the Chernoff bound value is just as good.